OM nucleic - nucleic search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: December 27, 2001, 20:39:15 ; Search time 7648.39 Seconds
               (without alignments)
               5864.742 Million cell updates/sec

Title: US-09-830-647-4
Perfect score: 2719
Sequence: 1 aattcggcacgagctctctg.........aaaaaaaaaaaaaactcgag 2719

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824858755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_vl:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2719	100.0	2719 2719	9	AB028070	Homo sapi
2	2648	97.4	2780	9	AB028069	Homo sapi
3	2326.8	85.6	2474	9	AF160249	Homo sapi
4	2324	85.5	2460	9	AF160876	Homo sapi
5	1986.6	73.1	236236	2	AC018977	Homo sapi
6	1427.4	52.5	2458	10	AF292400	Cricetulu
7	1194.2	43.9	122331	9	AC005164	Homo sapi
8	1194.2	43.9	163031	10	AC024947	Homo sapi
9	1194.2	43.9	163889	2	AC004958	Homo sapi
10	1129.6	41.5	2309	10	MMDBF4	Homo sapi
11	1101.6	40.5	2276	9	AK022969	Homo sapi
12	900.6	33.1	232903	2	AC091341	Homo sapi
13	554.6	20.4	136823	9	AC003083	Homo sapi
14	513.8	18.9	685	9	AX070167	Homo sapi
15	423.6	15.6	177899	2	AC074175	Mus muscu
16	414.4	15.1	433	11	G36582	SHGC-53468
17	411	15.1	422	6	AX072042	Sequence
18	384.4	14.1	403	6	AX070196	Sequence
19	313.2	11.5	205662	9	AC017002	Homo sapi
20	304.4	11.2	136555	2	AC068279	Homo sapi
21	304.4	11.2	157063	2	AC083899	Homo sapi
22	217.4	8.0	220	11	G41942	SHGC-64833
23	194.4	7.1	397	6	AX070116	Sequence
24	126.2	4.6	168717	2	AC024683	Homo sapi
25	126.2	4.6	177720	9	AC007514	Homo sapi
26	104.6	3.8	1517	9	AF125531	Homo sapi
27	99.4	3.7	377	6	AX070662	Sequence
28	98.6	3.6	7218	6	I66494	Sequence 14
29	67.8	2.5	1141	6	AX083744	Sequence
30	65.6	2.4	1141	6	AX083744	Homo sapi
31	64.2	2.4	87256	2	AC013754	Homo sapi
32	63.6	2.3	207420	2	AC078884	Homo sapi
33	63.4	2.3	321003	2	PFMAL4P3	Plasmodiu
34	61.2	2.3	71559	2	AC091013	Homo sapi
35	61	2.2	245802	9	AC006279	Plasmodiu
36	60.8	2.2	51274	2	AC023454	Homo sapi
37	60.8	2.2	167968	2	AC084398	Homo sapi
38	60.6	2.2	140176	2	AC016411	Homo sapi
39	60.6	2.2	174133	8	AF083031	Guillardi
40	60.2	2.2	201676	2	AC090647	Mus muscu
41	59.8	2.2	120091	2	AC012684	Homo sapi
42	59.8	2.2	170102	9	AC008079	Homo sapi
43	59.6	2.2	101109	2	AC027353	Homo sapi
44	59.4	2.2	62164	2	AC068512	Homo sapi
45	59.4	2.2	158167	2	AC034105	Homo sapi

ALIGNMENTS

RESULT 1
AB028070
LOCUS   AB028070   2719 bp   mRNA   PRI   04-JUN-1999
DEFINITION Homo sapiens mRNA for activator of S phase kinase, complete cds.
ACCESSION   AB028070
VERSION   AB028070.1 GI:4996097
KEYWORDS   activator of S phase kinase.
SOURCE   Homo sapiens HeLa cDNA to mRNA, clone:H37.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS   Kumagai,H., Sato,N., Yamada,M., Mahony,D., Seghezzi,W., Lees,E.,
            Arai,K. and Masai,H.
TITLE   A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates
            Human Cdc7-Related Kinase and Is Essential for G1/S transition in
            Mammalian Cells
JOURNAL   Mol. Cell. Biol. (1999) In press
REFERENCE   2 (bases 1 to 2719)
AUTHORS   Masai,H. and Kumagai,H.

8

```
TITLE       Direct Submission
JOURNAL     Submitted (28-MAY-1999) to the DDBJ/EMBL/GenBank databases. Hisao
            Masai, Institute of Medical Science, University of Tokyo,
            Department of Developmental Biology; 4-6-1 Shirokanedai, Minato-ku,
            Tokyo 108-8639, Japan (E-mail:hisao@ims.u-tokyo.ac.jp,
            Tel:81-3-5449-5661, Fax:81-3-5449-5424)
FEATURES             Location/Qualifiers
     source          1..2719
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="HeLa"
                     /chromosome="7"
                     /clone="H37"
                     /map="7q21"
     gene            518..1222
                     /gene="ASK/ H37 (huDbf4)"
     CDS             518..1222
                     /gene="ASK/ H37 (huDbf4)"
                     /note="C-terminally truncated due to alternative splicing;
                     an alternative splicing variant of ASK, a regulatory
                     subunit for hucdc7 kinase. This protein could bind to
                     hucdc7 but may not be able to activate its kinase activity."
                     /codon_start=1
                     /product="activator of S phase kinase"
                     /protein_id="BAA78327.1"
                     /db_xref="GI:4996098"
                     /translation="MNSGAMRIHSKGHFPQGGIQVKNEKNRPSLKSLKTDNRPEKSKCK
                     PlWGKVFYLDLPSVTISEKLQKDIKDLGRVEEFLSKDiISYILISNSKKEAKFAQTLGRI
                     SPVPSPESAVTAETGSPHPSHDGSSFKSPSDTVCLSRGKLLVEKAIKDHDFIPSNILS
                     NALSWGVKILHIDIRYYIEQKHKELYLLKKSSTSVRDGGKRVGSGAQKTRTGRLKKP
                     FVKVEDMSQSPAVIHLM"
     polyA_site      2563..2569
BASE COUNT      932 a    522 c    607 g    658 t
ORIGIN
```

```
Query Match   100.0%;  Score 2719;  DB 9;  Length 2719;
Best Local Similarity   100.0%;
Matches 2719;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
                                  Pred. No. 0;
```

```
Qy      1 aattcggcacgagagctcttctgaggctgcgccaagaacctgaagcggcggagagccggag 60
          ||||||||||||||||||||||| ||||||| |||||||||||||||||||||||||||||
Db      1 AATTCGGCACGAGAGCTCTTGAGGCTGCCAAGAACCTGAAGCGGCGGAGAGCCGGAG 60

Qy     61 gtctgagactgagagagacgcaaacggaatggaagcggcgggaaacacaactgcag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTCTGAGACTGAGAGAGACGCAAACGGAATGGAAGCGGCGGGAAACACAACTGCAG 120

Qy    121 ggcagagaggcgagcggccagaagaaggacgcgcagcgtgaaggggcgcgcagcgggagaa 180
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGCCAGAGAGGCGAGCGGCCAGAAGAAGGACGCGCAGCGTGAAGGGGCGCGCAGCGGGAGAA 180

Qy    181 ggcaggcacgaggggcgagcgcgaggcgggggcacgcgctggcgttgagacgggcgcgggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCAGGCACGAGGGGCGAGCGCGAGGCGGGGGCACGCGCTGGCGTTGAGACGGGCGCGGGG 240

Qy    241 cgcgcgtatcggcgcgggcgggccgtgacgcgtttcaaaatcttcaacgccgcagccca 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGCGCGTATCGGCGCGGGCGGGCCGTGACGCGTTTCAAAATCTTCAACGCCGCAGCCCA 300

Qy    301 ctcgtttgtcttgcgccctcctcccgcgccttggagcggaggccccccgaaac 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTCGTTTGTGTTGCTTTGCGCCCTTCCTCCCGCGGGATCCGGAGCCCCCGGAAAC 360

Qy    361 ccgaactgcagacgcggtaccctcctactcgctgactcggtcggagaggagagagg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCGAACTGCAGACGCGGGTACCTCTACTCGCTGACTCGGTCGGAGAGGAGAGAGG 420

Qy    421 cggcgtcctgtcaacagccgggggaagccgtgctttcgcgcgtcgcccgtgcacact 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CGGCCGTCGTGTCAACAGCCGGGGAAGCCGTGCTTTCGCGCGTCGCCCGTGCCGACACT 480

Qy    481 tcctccggacccagcaatgtaggtcgcgggactgccatgaactccggagccatgaggat 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCTCCGGACCCAGCAATGTAGGTCGCGGGACTGCCATGAACTCCGGAGCCATGAGGAT 540
```

```
Db    541 CCACAGTAAAGGACCCAGCATGTAGGTGCCGGGGACGTGCCATGAACTCCGGAGAGGAT 600
Qy    541 ccacagtaaaggaccattccagggtggatctccaagtcaaaaaatgaaaaaaaagaccatc 600

Qy    601 tctgaaatctctgaaaactgataaacaggccagaagaaatccaaagtaagccactttgggg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTGAAATCTCTGAAAACTGATAAACAGGCCAGAAGAAATCCAAAGTAAGCCACTTTGGGG 660

Qy    661 aaaagtatttaccttgacttacctcctgtcaccacatatctgaaaaacttcaaaggacat 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AAAAGTATTTTACCTTGACTTACCTTCTGTCACCACATATCTGAAAAACTTCAAAGGACAT 720

Db    721 TCTGAAATCTCTGAAAACTGATAAACAGGCCAGAAGAAATTCTCAGCAAAGATATCAGTTATTTT 780
Qy    721 taaggatctgggagggggcgagttcagaccaagaatttctcagcaaagatatcagttatctttc 780

Qy    781 aaataagaggaagctaaaatttgcacacaaaccttgggtgaatttctcctgtaccaagtcc 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AAATAAGAGGAAGCTAAAATTTGCACACAAACCTTGGGTGAATTTCTCCTGTACCAAGTCC 840

Db    841 AGAATCTGCATAATACTCCAGAAACCACTTCACCCYCATCCCAGCCATCCATGGAAGTTCATT 900
Qy    841 agaacctgcatatactgcagaaaccacttcaccccatcccagccatggaagttcatt 900

Qy    901 taagtcaccagacacagtgtttaagcagagagagaaaattattagtgaaaaagctatcaa 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TAAGTCACCAGACACAGTGTGTTTAACCAGAGGAAAATTATTAGTTGAAAAAGCTATCAA 960

Qy    961 ggaccatgatttattcctcaaatagtatattatcaaatgccttgtcatggggggtaaa 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGACCATGATTTTATTCCTTCAAATAGTATATTATCAAATGCCTTGTCATGGGGACTAAA 1020

Qy   1021 aattcttcatattcaatgacaattagatactacaattgaacaaaaaaaaaggttgtattt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AATTCTTCATATTGAATGACAATTAGATACTACAATTGAACAAAAAAAAAGAGTTGTATTT 1080

Db   1081 ACTCAACAAAAATCAAGTACTTCAGTAAGGAGGAGAGAAAAGAGTTGCTAGTGGTGCACA 1140
Qy   1081 actcaagaaaatcaagtacttcagtaaggagggcaaaagacttggtgagatgtaagcaaag 1140

Qy   1141 aaaaacaagaacagaggaagagactcaaaaagcctcttttgtaaagtggaagatataaga 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AAAAACAAGAACAGAGGAAGAGACTCAAAAAAGCCTTTTGTAAAGGTGGAAGATATAAGA 1200

Qy   1201 ccctgcagtccattgatgtagacaaagccacctctagtcaaagcaaagctaaggttaaa 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCCTGCAGTCCATTTGATGTAGACAAAGCCATCTAGTAGTCAAAGCAAAACTCAGGGTTAAA 1260

Db   1261 CTAAGAAATCCAAACAGCATGGCGATAAGTTAGTGGAACCTCAATTCAACTCCAGTTCAAA 1320
Qy   1261 ctaagaaatccaaacagatggcgataagtatgtggaacctcaattcaactccagtgaaa 1320

Qy   1321 gagaagaagaaaaaggtatattgtttcctcagaaaattattgaacatcagaaact 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GAGAAGAAGAAAAAGGTATTTGTTTCCTTCAGAAAATATTGAAACATCTAGAAACT 1380

Db   1381 CACCTTCTAAGTGAGCAACAAACTTTGCACAGAAGTAACCAGTATCAAGTTGTTGCAT 1440
Qy   1381 cacttctaagtgagcaacaacagacatttgacagaagaaaacttgtcaaagttgtgat 1440

Qy   1441 gatattgtatctcaagttagttttctgaaatatgaaaggaaacacacctaaaaag 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GATATTGTATCTCAAGTTAGTTTTGACTTTTGTGGAATATGAAAGGAACACACCTAAAAAG 1500

Db   1501 AAAAGAATAAAATACAAGTGTTTGGAATCCCTTTCTCTCCGTTTTCTGCAAGTGTTCCTGAAAAAG 1560
Qy   1501 aaaagaataaaatacaagtgtttgggatcccttttctctccctgttctctgaaaag 1560

Qy   1561 actgacaaaggagaaaatgtggaattgcaacatattctcagaaagagattgccaggaaggat 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACTGACAAAGGAGAAAATGTGGAATTGCAACATATTCTCAGAAAGAGATTGCCAGGAAGGAT 1620
```

```
Db  1561 ACTGNAACAAAAGGAAAAAGTGGAATTGCAACACATATTTCTCAGAAGATTGCCAGGAAGAT 1620

Qy  1621 gatacaaacagtgaggagagcagagaattttcctgtataaagagagaacctgaaaaaag 1680
         |||||||||||||| ||||| ||||||||||||| |||||||||||||| |||||||||
Db  1621 GATACAAACAGTGAAGAGGAGCAGAGAATTTCCTGTATAAAGAGAGAAACTGAAAAAAG 1680

Qy  1681 ctcctgttattcaagagcccatcccccaccctcaaatgaattgagagggcttaatgag 1740
         |||| ||| |||| ||||||||||||||||| |||||||||||||||||||||||||||
Db  1681 CTCCCTGTTTATTTCAGAGCCCATCCCCCACCCTTCAAATGAATTGAGAGGGCTTAATGAG 1740

Qy  1741 aaaatgagtaataaatgttccatgttaagtacacgtgaagagaattttt 1800
         |||||||||||||| |||||||||||||||||||||| |||||||
Db  1741 AAAATGAGTAATAAAATGTTCCATGTTAAGTACACGCTGAAGAGAACACCGAAAAAAAG 1800

Qy  1801 acacagctacctctacatataaaaacacaggaatgcaattcttgacattccgaacacaca 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 ACACAGCTACCTCTACATATAAAAACACAGGAATGCAATTCTTGACATTCCGAACACACA 1860

Qy  1861 ttaagtgaaaatgacttagaaagaaactaagggtagaactcacatacataacacaggca 1920
         |||| |||||||||||||||| |||||||||||||||| ||||| |||||||||||||||
Db  1861 TTAACTGAAAATGACTTAGAAGAACTAAGGGTAGATCACTATAAAAGTAACATACAGGCA 1920

Qy  1921 tctgtacatgttctgattcagtgaatctcaaccaaaaaacagaagtca 1980
         |||||||||||| |||||||||||||||||||||||||||||||||||
Db  1921 TCTGTACATGTTTCTGATTCAGTGAATCTCAACCAAAAAACAGAAGTCA 1980

Qy  1981 gatatgtgctttttccaggaaggggaaagaggactcaatattactact 2040
         |||| ||||||||| ||||||||| ||||         |||| |||
Db  1981 GATACTGTGCTTTTTCCAGGAAGGGAAACCCAATGAATATATTTACT 2040

Qy  2041 catgattctcgtctgataacaataaacagttcacaagacaactgttcagggcaaag 2100
         |||||||||||||||||||||||||||||||||||| ||||||||||||| |||||
Db  2041 CATGATTCTCGTCTGATAACAATAAACAGTTCACAAGAGCAACTGTTCAGGCAAAG 2100

Qy  2101 gctccatccatcctcctgaggaaccccaatgaattgacttcaagaatatggatagt 2160
         |||||| |||| |||||||||||||||||||||||   |||||||||||||||||
Db  2101 GCTCCATTCCATTCCTCCTGAGGAACCCCAATGAATTCAAGAATATGGATAGT 2160

Qy  2161 ttaccttctggttaaatacatccgaaaatgaaataatattaggacgaacaataagaaaa 2220
         |||||||||| |||||||||| |||||| || |||||||||||||| ||| |||||||
Db  2161 TTACCTTCTGGTAAAATACATCGAAAAATAATTAATTAGGACGAAAATAGAAAAGAA 2220

Qy  2221 aatctgaaccaaatgctgaaattgataaaagaaactgaattattaccacagaagaaaac 2280
         ||||||||||||||||||||||| || ||   |||| |||||||| |||||||||||||
Db  2221 AATCTGAACCAAATGCTGAAATTGTGAATTTATTAAAGACAACTAATTATTAACACAAGAGAACAAAAC 2280

Qy  2281 agaaattgtagttcacggtacagtcttttcagaagatgttgaagagaaa 2340
         |||||||||||||||| |||||||||| |  |||  ||| |||||||||
Db  2281 AGAAATTGTAGTTCACCGGTACAGTCTTTACTAGACTAGTGAAGAGAAA 2340

Qy  2341 tcagaaatttttgggtttcacaagctacaccatgttttagat 2400
         |||||| ||||||||||| ||||||||||||   |||| |||
Db  2341 TCAGAATTTTTGGGTTTTCACAAGCTACACACAGGAAATTAGAAATAGAAAGAT 2400

Qy  2401 atttggaggagagaaattcagatatctgttaacacgcgttttctccttcaact 2460
         ||||||| ||| |||||||||||| || |||||||| ||||| |  |||||||
Db  2401 ATTTGGAGCAGGAAAATTCAGATAATCGTTAACACGGTTTTTCTGGTCCCTTCAACT 2460

Qy  2461 tctacaattactggctttagaaattgatacactttgagaggatacggatc 2520
         ||||| |||||||||||| |||||| |||||| ||| ||||| |||||||
Db  2461 TCTACATTTACTGGCTTTTAGAATTGCAATACTTTTCAGAAGTGATAAGGATC 2520

Qy  2521 atattccttgaaattttatataaatatgtggaaattcttaggagatttcttaccagcgtttg 2580
         |||||| |||||||||| ||||||||||| |||||||||||   |||||| |||||| ||||
Db  2521 ATATTCTTGAAATTTTTATAAATATGTATGGAAATTCTTAGGAATTTTTTACCAGCTTTTG 2580

Qy  2581 tttacagacgcccaatgtaaatattaaaaataaatatttgcaattctacagaattgaat 2640
         |||||||||| ||||||||||| |||||||||||||||||||||| ||||||||||||
Db  2581 TTTACAGACCCAAATGTAAAAGTAAAATATTAAAAATAAATATTTGCAATTTCTACAGAATTGAAT 2640

Qy  2641 acctgttaaagaaaaattacaagataaattacttgactggtcttgtttctacagaattg 2700
         ||||||||||||||||||||| ||| ||| |||| |||||||| |||||| || |||
Db  2641 ACCTGTTAAAGAAAAATTACAGAATAAAAACTTGTGACTGGTCTGTTTTTACATTAAAAAA 2700
```

RESULT 2
AB028069

```
Qy  2701 aaaaaaaaaaaaaaactcgag 2719
         ||||||||||||||   ||||
Db  2701 AAAAAAAAAAAAAAACTCGAG 2719
```

LOCUS   AB028069   2780 bp   mRNA   PRI   03-JUN-1999
DEFINITION Homo sapiens mRNA for activator of S phase kinase, complete cds.
ACCESSION   AB028069
VERSION   AB028069.1 GI:4996095
KEYWORDS   activator of S phase Kinase.
SOURCE   Homo sapiens Hela cDNA to mRNA, clone:H37.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
  AUTHORS   Kumagai,H., Sato,N., Yamada,M., Mahony,D., Seghezzi,W., Lees,E.,
            Arai,K. and Masai,H.
  TITLE   A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates
            Human Cdc7-Related Kinase and is Essential for G1/S transition in
            Mammalian Cells
  JOURNAL   Mol. Cell. Biol. (1999) In press
REFERENCE   2 (bases 1 to 2780)
  AUTHORS   Masai,H. and Kumagai,H.
  TITLE   Direct Submission
  JOURNAL   Submitted (28-MAY-1999) to the DDBJ/EMBL/GenBank databases. Hisao
            Masai, Institute of Medical Science, University of Tokyo,
            Department of Developmental Biology; 4-6-1 Shirokanedai, Minato-ku,
            Tokyo 108-8639, Japan (E-mail:hisao@ims.u-tokyo.ac.jp,
            Tel:81-3-5449-5661, Fax:81-3-5449-5424)
FEATURES   Location/Qualifiers
     source   1..2780
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="Hela"
                     /chromosome="7"
                     /clone="H37"
                     /map="7q21"
     gene   518..2542
                     /gene="ASK/H37 (hudbf4)"
     CDS   518..2542
                     /gene="ASK/H37 (hudbf4)"
                     /function="binds and activates hucdc7 kinase"
                     /note="ASK encodes a regulatory subunit for hucdc7, the
                     human homologue of budding yeast Cdc7 kinase.; Amino acid
                     51-90 and 290-322 are motif-N and motif-C, respectively,
                     conserved in Dbf4-related molecules. Amino acid 201-218 is
                     a putative bipartite nuclear localization signal. Amino
                     acid 101-120 and 552-564 are two putative PEST-like
                     sequences."
                     /codon_start=1
                     /product="activator of S phase Kinase"
                     /protein_id="BAA78326.1"
                     /db_xref="GI:4996096"
                     /translation="MNSGAMRIHSKGHFQGGIQVKNEKNRPSLKSLKTDNRPERSKCK
                     PLWGKVFYLLDPSVTISEKLQKDIKDLGGRVEEFLSRDISYLISNKREAKFAQTLGRI
                     SPVPSPESAYTAETTSPHPSHDGSSFRSKSPDTVCLSRGKLLVRKAIKDHDFIPSNILS
                     NALSWGYKILHIDDIRYYIEQKKELYLLKSSTSVRDGGKRVGSGAQKTRVGRLKKP
                     FVKVEDMSQLYRPFYLQITMPFINYSTQKPCSPFDVDKPSSMQKQTQVKLRIQTDGD
                     KYGGTSIQIOLKEKKKGYCECCQLKYEDLETHILSEQHRNFAQSNQYQVVDDIVSKL
                     VEFDFVEVEKDTPKKKRIKYSVGSSLSPVSASVLKKTEQKEKVELQHISQKDCQEDDTYV
                     KEQNFLYKETQEFEKKLLFISEPIPHPSNELRGLNERKMSNKCSMLSTAEDDIRQNFTQ
                     LPLHKNKQECILDISEHTLSENDLEELRVDHYKCNIQASVHVSDFSTDNSGSQPKQRS
                     DTVLFPAKDLKEKDLHSIFTHDSGLIFINSSQEHLTVQAKAFHTPPEPNECDFKNM
                     DSLPSGKIHRKVKIILGRNRKENIEPNAEFDKRTEFITQEENRICSSPVQSLLDLFQT
                     SEEKSEFLGFISYTEKSGICNVLDIWEEENSDNLLTAFFSSPSTSFTGF"
     polyA_signal   2724..2730
BASE COUNT   949 a   534 c   613 g   684 t
ORIGIN

Query Match   97.4%;   Score 2648;   DB 9;   Length 2780;
Best Local Similarity   97.8%;   Pred. No. 0;
Matches 2719;   Conservative   0;   Mismatches   0;   Indels   61;   Gaps   1;

```
Qy     1 aatcggcacgagctctctgagtcgtcgcccaagacctgaagcggagagcccgg 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AATTCGGCACGAGCTCTCTCAGGCTGCCCAAGACCTGAACGCGGGAGCCCGG 60

Qy    61 gtctgagactgagagagacaacggaatgagagcggggtagagggcggaaacacacacctgcag 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GTCTGAGACTGAGAGAGACAACGGAATGGAGGCGGGGTAGAGGCCGGAAACACACCTGCAG 120

Qy   121 cggcgagagccggaggcgcgcggagaaaggaccggcgccgcgtgaggggggccgagcgccgagcgagaa 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGCCGAGCGGCAGGCGCGCAGGAGGACGGCGCCGCGTGAGGGGGGGCGGGCGGGGGCCGGGGCCGGAGGA 180

Qy   181 ggcaggcacgaggggggagagcgcgagggggcacggggggccgtcgagacgcgggggcggg 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCAGGCACGAGGGGGGGAGAGCGGCGGAGGGGGCACGGGGGGCCGTGGCGTGAGACGGGCGGGG 240

Qy   241 cggcggtatcgcgcgcccgcgcgtgacgcgtgtttcaaaatcttcaaccgcagcccca 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CGGCCGTATCGCGCGGCCGCGTCAACAGCGCGGGAAACGCCGTTTTTCAAAATCTTCAACTCCACTGCACC 300

Qy   301 ctcgttttgtctttgcgccttcctcctcccgcgccctggagcggccggttcccgcgccggaaac 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTCGTTTGTGTTGCTTTGCCCGCTTCCTCTCCCGCGCCGGCGTTGGAGCGGCCGGGATCCGGCCCGGAAAC 360

Qy   361 ccgacctgcagacgcggtacctcactactggtaacaggccgagaaaatccaaatgtaagccacttggg 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCGACCTGCAGACGCGGTACCTCTACTGCGTAGAGGCCGAGAAGGAGAGAGG 420

Qy   421 cggcgtcctgtcaacaggccggggggaagccgctgtgcttcgcggctgccccgtgtcgacact 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CGGCCTCCTGTCAACAGGCGGTGCGGCCTGCCCCGGTGCCGGCTGCGGCGACACT 480

Qy   481 ttcctccggacccagtccatgtaggtccgggcgactgccatgaacctccggagccatcctaggat 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TTCTCCGGACCCAGCATGTAGGTGCCGGGGACTGCCCATGAACCTCCGGAGCCATCGAGGAT 540

Qy   541 ccacagtaaaggacacattcccagggtggaatccaagtcaaaaatgaaaaaacagaccccatc 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCACACTAAAGGACACATTTCCAGGGTGGAATCCAAGTCAAAATGAAAAAACAGACCCATC 600

Qy   601 tctgaaatcctgaaaaactgataacaggccagaaaatccaaatgtaagccacttggggg 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TCTGAAATCTCTGAAAAACTGATAACAGGCCAGAAAAATCCAAATGTAAGCCACTTTGGGG 660

Qy   661 aaaagtattttaccttgacttacctttcctgtcaccaatatctgaaaaaacttcaaaaggacat 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AAAAGTATTTTACCTTGACTTACCTTTCCTGTCACCAATATCTGAAAAACTTCAAAAGGACAT 720

Qy   721 taaggatctgggaggggagttgaagaatttctcagcaaagatcgttatcttatttc 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 TAAGGATCTGGGAGGGGAGTTGAAGAATTTCTCAGCAAAGATCACTTAGTTCATTTC 780

Qy   781 aaaataagaagagaagctaaaattgcacaaaaccttgggtcgaatttctcctcgtaccaagtcc 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AAATAAGAAGGAGAAGCTAAATTGCACAAAACCTTGGGTCGAATTTCTCCTGTACCAAGTCC 840

Qy   841 agaatctgcatatactgcagaaacacactccacctcaactcttagttgaaaagctatcaa 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AGAATCTGCATATACTGCAGAAACACCTCAACCTCAACTCTCAATGGAAGGCTCATT 900

Qy   901 taagtcaccagacacaacagtgtgttgttaagcgagaggaaaaattattagttgaaagtcatt 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TAAGTCACCAGACACAGTGTTTTAGCGAGGAAAAATTATTACTTGAAAAGCTATCAA 960

Qy   961 ggaccatgattttattcctcaaatagtatattccaatgccttgtcatgggagtaaa 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGACCATGATTTTTATTCCTTCAAATAGTATATTACCAGAGAAGGATCTCAAGGAAAAGGATCTCAAGTATATTTAC 1020

Qy  1021 aattcttcatattgatgacattagatacatcattgaacaaagaaaagagagagttgtatt 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AATTCTTCATATTGATGATGGATTGACATTAGATACTACATTGAACACAAAAAGAGAGAGTTGTATTT 1080

Qy  1081 actcaagaaatccaagtacttcagtaagatcgggggcaaaagactcaaggttaa 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACTCAAGAAAATCAAGTACTTCAGTAAGATGGGGGGCAAAAGAGTTGGGTGTAGTGGTTGCACA 1140

Qy  1141 aaaaacaagaacaggagggagagactcaaaaagcccttttgtaaagagtggaagataagca--- 1197
         |||||||||||||||||||||||||||||||||||||||||||||||||||| --
Db  1141 AAAAACAAGAACAGGAGGGAGAGACTCAAAAGCCTTTTTGTAAAGGTGGACAATATGAGCCAACT 1200

Qy  1198 --- 1199
         ||
Db  1201 TTATAGGCCATTTTATTCTTCAGCTGACCAATATGCCTTTTATAAAAATTATTCTATTCAGAA 1260

Qy  1201 gcctgcagtccatctttgatgtagacaagcgccatctagtatgtcaaatgaaactccagttaa 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GCCCTGCAGTCCATTTGATGTAGACAGCAGCCATCTAGTATGTCAAAATGAAACTCCAGTTAA 1320

Qy  1261 actaagaatccaacacagatgggcgataagtatggttggaacctcaattcaactccagttgaa 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ACTAAGAATCCAACAACAGATGGGCGATAAGTAAGTATGGTTGGAACCTCAATTCAACTCCAGTTGAA 1380

Qy  1321 agaagaggaagaaagaggatattgtaaatgttgcttgcagaaatatgaagtctagaaac 1379
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 AGNAGAGAAGAAAAGGAGAATTATTGTGAATGTTGCTTGCAGAAATATTAGAAAC 1440

Qy  1380 tcaccttctaagtgagcaacacacagagtaggtcgacagtatatcaagttgtttga 1439
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 TCACCTTCTAAGTGAGCAACACAGAAACTTCGACAGTATCAAGTTGTTTGA 1500

Qy  1440 tgatattgtatctaagttagtttttgtggaatatgaatagaggacacacttaaaa 1499
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 TGATATTGTATCTAAGTTAGTTTTTGTGGAATATGAAAGGAGACACACCTAAAAA 1560

Qy  1500 gaaaagaataaacacagtgttgggatccctttcctcctgttcctgaaaaa 1559
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GAAAAGAATAAAATAAACAGTGTTGGGATCCCTTTCTCCTGTTCCTGAAAAA 1620

Qy  1560 gactgaacaaaagtgggaattgcaacatattctcaggaagaagaa 1619
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 GACTGAACAAAAGTGGGAATTGCAACATATTTCTCAGGAAGAAGAGA 1680

Qy  1620 tgatacaaacagtgaaggagcagaatttcctgtatataagagaccccagaaactgaaaaaaa 1679
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGATACAACAGTGAAGGAGCAGAATTTCCTGTATAAAGAGACCCAGGAAACTGAAAAAAA 1740

Qy  1680 gtccctgtttatttcagaagcccatccccacccctccaaatgaattgatgagggcttaatga 1739
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 GCTCCTGTTTATTTCAGAAGCCCATCCCCACCCTCAAATGAATTGAGAGGGCTTAATGA 1800

Qy  1740 gaaaatgagtaataaatgttccatgttaagtacagctgaagtggcataagacagaatt 1799
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 GAAAATGAGTAATAAATGTTCCATGTTAAGTACAGCTGAAGTGGCATAAGACAGAATT 1860

Qy  1800 tacacagctacctctacataaaaaccaacggaatgcattcttgacattccgaacacac 1859
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TACACAGCTACCTCTCACATAAAAAACCAACGGAATGCATTCTTGACATTCCGAACACAC 1920

Qy  1860 attaagtgaaaatgacttagaagaaactaaggtagtcacttatcaatacagggc 1919
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ATTAAGTGAAAATGACTTAGAAGAACTAAGGGTAGATCACTTATCAATACAGGGC 1980

Qy  1920 atctgtacaatgttctgattctgacacagaataatagtgatctcaacacaaacaagaagtc 1979
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ATCTGTACAATGTTCTGATTTCAGATTTCACAGAATAAATAGTGAACTCAACAAAAGAGAAGTC 2040

Qy  1980 agatactgtgctttttcccacgcaaagggaaaggaagggacctcaattccaatatttac 2039
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 AGATACTGTGCTTTTTCCCACGGAAAGGGAAAGGGAAGGGACCTCAACCAAAAGCAGAAGTC 2100

Qy  2040 tcatgtacctcggtctgataacaataaaacagttcacaagacacctaactgttcaggcaa 2099
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT  3
LOCUS       AF160249     2474 bp    mRNA              PRI       01-JUL-1999
DEFINITION  Homo sapiens Dbf4 protein mRNA, complete cds.
ACCESSION   AF160249
VERSION     AF160249.1  GI:5306124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2474)
  AUTHORS   Jiang,W. and Hunter,T.
  TITLE     Mammalian Cdc7/Dbf4 Protein Kinase Complex is Essential for
            Initiation of DNA Replication
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2474)
  AUTHORS   Jiang,W. and Hunter,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Molecular Biology and Virology Laboratory,
            The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA
            9037, USA
FEATURES             Location/Qualifiers
     source          1..2474
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
     CDS             217..2241
                     /note="HsCdc7 protein kinase regulatory subunit"
                     /codon_start=1
                     /product="Dbf4 protein"
                     /protein_id="AAD41911.1"
                     /db_xref="GI:5306125"
                     /translation="MNSGAMRIHSKHIFQGGIQVKREKNRPSLKSLKTDNRPEKSKCK
                     PLWGKVFYLLDLPSVIISEKLQDGVEEFLSPDTVCLSRGKLLVEKAIKDHFIPSNSILS
                     SPVPSPESAYTAETTSPHDSHDGSSFKSPDTVCLSRGKLLVEKAIKDHFIPSNSILS
                     NALSWCKVKILIHDDIRYYIEQKKKELYLLKKSSSTSVRDGVRGVGSAGKTRTGRLKKP
                     FVKVEDMSQLYRPFYLQLTNMPFINYSIQKPCSPFDVDKPSSMQKQTQVKLRIQTDGD
                     KYGGTSIQLQKEKKKGYCECCLQKYEDLETHLLSEQHRNFAQSNQYQVVDDIVSKL
                     VFDFVEYEKDPKKKIKYSVGSSLSPVSASVLKKTEQKEVKELQHISQKDCQEDDTTV
                     KEQNFLYKETQETEKKILFISEPIPHPSNELRGLNEKMSNKCSMLSTAEDIRQNFTQ
                     LPLHKNNQEECLLDISEHTLSENDLEELRVDHYKCNIQASVRHVSDFSTDNSGSQPKQKS
                     DTVLFPAKDLKEKDLHSIFTHDSGLTINSSQEHLTVQAAKAPFHTPEEPNECDFRNM
                     DSLPSGKIHRVKIILGRNRKENLEPNAEFDKRTEFITQENRICSSPVQSLLDLFQT
                     SEEKSEFLGFTSYTEKSGICNVLDIWEEENSDNLLTAFFSPSTSTPTGF"
BASE COUNT      888 a    449 c    480 g    657 t
ORIGIN
```

```
Query Match            85.6%;  Score 2326.8;  DB 9;  Length 2474;
Best Local Similarity  97.4%;
Matches 2410;  Conservative  0;  Mismatches  2;  Indels  62;  Gaps  2;

Qy    301 ctcgtttgtgcttttgcgccttcctccctcccggggatcccggccttggagcggatccggaaac 360
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db      1 CTCGTTTGTGCTTTTGCG-CTTCCTCCCTCCCGGGCCCTTGGAGCCGGCGCGCCCGGAAAC 59

Qy    361 ccgacctgcagacgcgggtacctctactgcgtaggaggccgtagctggcggaaggagagagg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 CCGACCTGCAGACGCGGGTACCTCTACTGCGTAGCTGGCGGAAGGAGAGAGG 119

Qy    421 cggccgtcctgtcaaaccaggccgggggaagccgtgcttctcgcggctgccggtggcgacct 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 CGGCCGTCGTCAACAGGCCGGGGGAAGCCGTGCTTCTCGCGGCTGCCGGTGGCGACACT 179

Qy    481 ttctccggaccagccatgtaggtgccgggccgactgaactgccatgagctccggaccggat 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 TTCTCCGGACCCAGCCATGTAGGTTGCCCGGGGCGGGACTGCCATGAGCTCCGGAGGAT 239

Qy    541 ccacagtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacgaaccctc 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CCACAGTAAAGGACATTTCCAGGGTGGAATCCAAGTCAAAAATGAAAAAAACAGAACATC 299

Qy    601 tctgaaatcctctgaaaactgataaacaggcccagaaaaatcccaaatgtaagcccactt 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 TCTGAAATCTCTGAAAACTGATAAACAGGCCCAGAAAAATGTAAGCCACTTTGGGG 359

Qy    661 aaaaagtatttacttgacttacttctgtcaccatatctgaaaaacttcaaaaggacat 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 AAAATAAGAAAGGAAGCTAAAATTTGCACAAAACCTTGGGTCAACATTCTCCTGTACCAAGTCC 419

Qy    721 taaggatactgtggagggccgagttgaagaattctccagcaaagatatcagttatcttattc 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 TAAGGACTCTGGGAGGGGCAGTTGAAGAATTTCTCAGCAAAGATATCAGTTATCTTATTTC 479

Qy    781 aaataaggaagctaaaatttgcacaaaaccttgggtcgaattcctcctgtaccaagtcc 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AAATAAGGAAGGAAGCTAAATTTGCACAAAACCTTGGGTCGAATTCTCCTGTACCAAGTCC 539

Qy    841 agaatctgcatactgcagaaccacttcacctcaccagccatggtggaagttcatt 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 AGAATCTGCATACTGCAGAAACCACTTCACCTCAATCGATGGAAGTTCATT 599

Qy    901 taagtcaccagacacagtgtgttttaagcagggaaaattaattagttgaaaagctatcaa 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 TAAGTCACCAGACACAGTGTGTTTTAAGCAGGGAAAATTATTAGTTGAAAAGCTATCAA 659

Qy    961 ggaccatgatttattccttcaaaatagtatattatccaaatgccttgtcatgggagtaaa 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 GGACCATGATTTTATTCCTTCAAAATAGTATATTATTCCAAATGCCTTGTCATGGGAGTAAA 719
```

```
Db   2100 ggctccattccataactcctcctgagggaacccaatgaatgtgactccaagaatagggatag 2159
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TCATGATTCTGTCTGATAACAATAAACAGTTCACAAGAGCACTTAACTGTTCAGGCAA 2160

Db   2161 GGCTCCATTCCAATCACTCCCTCGAGGAACCCAATGAATGAGACTTCAAGAATATGGATAG 2220
Qy   2160 GGCTCCATTCCAATCACTCCCTCGAGGAACCCAATGAATGAGACTTCAAGAATATGGATAG
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2220 ttaccttctggtaaaatacatcgaaaatgaaaatatattatgacgaaatagaaaaga 2219

Db   2221 TTTAACTTCTGGTAAAATACATCGAAAAAGTGAAAATATTAATATTAGGAGAAATAGAAAGA 2280

Qy   2280 aaatctggaaccaaatgctgaatttgataaagaaactgaattttattaccacaagaagaaa 2339
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AAAATCTGGAACCAAATGCTGAATTTGATAAAGAAACTGAATTTTATTACCACAAGAAGAAA 2340

Qy   2340 cagaatttgtagttcacccgtacagtctcttcagactagttgaagagaa 2399
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CAGAATTTGTAGTTCACCCGGTACAGTCTCTTTACTACGACTAGTTGAAGAGAA 2400

Qy   2400 atcagaaattttggcttcacagaacttcacagaagagagtggtatatgtcaatgttttaga 2459
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 ATCAGAAATTTTGGCTTCACAGAACTTCACAGAGGTGTATATGTCAATGTTTTAGA 2460

Qy   2460 ttctacattcactggcttttagaattcaatgcatactttttcaagaagtgataaggat 2519
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TATTTGGGAAGGAGGAAAATTCAGATAATAAATCTGTTAACACGGGTTTTTCTCGTCCCCTTCAAC 2520

Qy   2520 catattcttgaaatttttatatatatgtgtggaaattctttaggagttcttaccagcttt 2579
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TTCTACATTTACTGGCTTTTAGAATTTAAAATCCATACTACTTTTCAGAAGTGATAAGGAT 2580

Qy   2580 gttttacagacccaaatgtaaatcagacatatttgcaaattcctctacagagaattgaa 2639
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CATATTCTTGAAATTTTTATATATGTGTGGAAATCTTTAGGATTTTTTTACCACCTTT 2640

Qy   2640 tacctgttaaagaagaaataacacagaataaaacttgtgactggtcttgttgtttacattaaaaa 2699
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GTTTACAGACCCAAATGTAAATATTAAAATAATATTTGCAATTTCTACAGAAGAATTGAA 2700

Qy   2700 aaaaaaaaaaaactcgag 2719
          ||||||||||||||||||
Db   2701 TACCTGTTAAAGAAGAAAATTACAGAATAAAAACTTGTGACTGGTCTTGTTTACATTAAAAA 2760

Db   2761 AAAAAAAAAAAAACTCGAG 2780
```

```
Qy  1021 aatttcttcattcgattcaatcaattagagatactacttgagaacaacaagaaaaagaagaggttgtattt 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 AATTCTTCATTCGATTCAATGACATTAGACATTACATTGAACAAAAAAAGAGTTGTATTT 779

Qy  1081 actccaagaaatcaagtacttcagtaagagatgggggcaaaaagaagaggttggtagtggttgcaca 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 ACTCAAGAAAATCAAGTACTTCAGTAAGAGATGGGGGCAAAAAGAGTTGGTAGTGGTTGCACA 839

Qy  1141 aaaaacaagaacaggagaagactcaaaaagagcctttttgtaaaggtgaaagatatgagcca--- 1197
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   840 AAAAACAAGAACAGGAAGACTCAAAAAAGCCCTTTTGTAAAGGTGGAAGATATGAGCCAAACT 899

Qy  1198 ------------------------------------------------------------aa 1199
Db  1198 ------------------------------------------------------------aa 1199

Qy   900 TTTATAGGCCATTTATCTTCAGCTGTGACCAATATGCCTTTTATAAATTATTCTATTCAGAA 959
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 TTTATAGGCCATTTATCTTCAGCTGTGACCAATATGCCTTTTATAAATTATTCTATTCAGAA 1259

Qy  1200 gccctgcagtccattcgatgtagtacagcacctctagtatgtgaaaagacaaactcaggttaa 1259
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 GCCCTGCAGTCCATTTGATGTAGTACAGCACCATCTAGTATGTGAAAAGACAAACTCAGGTTA 1019

Qy   960 actaagaaatccaaacagagtggccgataagctataggtgaacctccaattcaactccagttgaa 1019
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260 ACTAAGAAATCCAAAACAGAGTGGCGATAAAGCTATAGGTGAACCTCAATTCAACTCCAGTTGAA 1319

Qy  1260 actaagaaatccaaacagagtggccgataagctataggtgaacctccaattcaactccagttgaa 1319
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 ACTAAGAAATCCAAAACAGAGTGGCGATAAAGCTATAGGTGAACCTCAATTCAACTCCAGTTGAA 1079

Qy  1020 tatttgggagaggagaaaattcaagtaaatctgttaacacgcgtcttttctcgtccccctccaac 1079
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2400 TATTTGGGAAGAGGAGAAAATTCAAGTAAATCTGTTAACACGCGTCTTTTTCTCGTCCCCTCAAC 2459

Qy  1320 agagaagaagaagaaaaggattattgtaatgttgcttgcgcagaaatatgaagattcagaaac 1379
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2160 TATTTGGGAAGAGGAGAAAATTCAAGTAAATCTGTTAACACGCGTCTTTTTCTCGTCCCCTCAAC 2219

Qy  1080 tcacttcttctaagtggcaacacacagaaacttcgcacagagtaacccagcatacaagttgttga 1139
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2460 TTCTACAGTTTAACTGCGGCTTTTACAGAATTTTATGGGAAGAATTACGATAAAAATGCAGTTTT 2519

Qy  1380 tcacttcttactgtggcaacacagaaacttcgcacagagtaacccagcatacaagttgttga 1439
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2220 TTCTACAGTTTACTGCGGCTTTTACAGAATTTTATGGGAAGAATTACGATAAAAATGCAGTTTT 2279

Qy  1140 TGATATTGGTATTCAAGTTAGTTTTTGACTTTGTGGAAATATGAAAAGGACACACCTAAAAA 1199
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2460 TTCTACAGTTTACTGCGGCTTTTACAGAATTTTATGGGAAGAATTACGATAAAAATGCAGTTTT 2519

Qy  1440 tgatttgtattctaagttagttttttgacttttgtgggaatatgaaaaggacacacctaaaa 1499
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2040 CAGAATTTGTTAGTTCACCGGGTACACAGTCTTTACTTGGACTTGTTCTTCAAGACGACTAGAGA 2099

Qy  1200 TGATATTGGTATTCAAGTTAGTTTTTGACTTTGTGGAAATATGAAAAGGACACACCTAAAAA 1259
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1200 TGATATTGGTATTCAAGTTAGTTTTTGACTTTGTGGAAATATGAAAAGGACACACCTAAAAA 1259

Qy  2340 atcagaattttgggttcacaagctaacacagaaaagaagagagagtggtatgtcaatgttttaga 2399
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2040 ATCAGAATTTTGGGTTCACAAGCTAACACAGAAAAGAAGAGAGTGGTATGTCAATGTTTTAGA 2159

Qy  1500 gaaaagaataaaatacaagtgtgttggatccccttctctcgttctgttctgcaagtgtcctgaaaa 1559
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1500 GAAAAGAATAAAATACAGTGTGTTGGATCCCCTTCTCTCGTTCTGTTCTGCAAGTGTCCTGAAAA 1559

Qy  2400 tatttgggaagaggagaaaattcaagtaaatctgttaacacgcgtcttttctcgtccccctccaac 2459
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2100 TATTTGGGAAGAGGAGAAAATTCAAGTAAATCTGTTAACACGCGTCTTTTTCTCGTCCCCTCAAC 2159

Qy  1260 gaaaagaataaaatacaagtgtgttggatccccttctctcgttctgttctgcaagtgtcctgaaaa 1319
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1260 GAAAAGAATAAAATACAGTGTGTTGGATCCCCTTCTCTCGTTCTGTTCTGCAAGTGTCCTGAAAA 1319

Qy  2460 ttctacagttttaactgcggcttttacagaattttatgggaagaattacgataaaaatgcagttt 2519
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2160 TATTTGGGAAGAGGAGAAAATTCAAGTAAATCTGTTAACACGCGTCTTTTTCTCGTCCCCTCAAC 2219

Qy  1560 gactgaacaaaaggagaattggaattgcaacatatttctcaagaaagaattgccaggaaga 1619
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1560 GACTGAACAAAAGGAGAATTGGAATTGCAACATATTTCTCAAGAAAGAATTGCCAGGAAGA 1619

Qy  2640 tacctgttaagagaaaattacagagaataaaacttgttgactggtctctgttcttacattaaaaa 2699
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2340 GTTTACAGACACCCAAATGTAAAATGTAAATTTGCAATTTTCTACAGAANTTGAA 2399

Qy  1320 gactgaacaaaaggagaattggaattgcaacatatttctcaagaaagaattgccaggaaga 1379
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1320 GACTGAACAAAAGGAGAATTGGAATTGCAACATATTTCTCAAGAAAGAATTGCCAGGAAGA 1379

Qy  2700 aaaaaaaaaaaaaaa 2713
         |||||||||||||||
Db  2460 AAAAAAAAAAAAAAA 2473
```

```
gene            1..2460
                /gene="DBF4"
                /map="7q21"
                /chromosome="7"
                /db_xref="taxon:9606"
CDS             200..2224
                /gene="DBF4"
                /product="DBF4-like protein"
                /codon_start=1
                /protein_id="AAD45357.1"
                /db_xref="GI:5566281"
                /translation="MNSGAMRIHSKGHFGQGGIQVKNEKNRPSLKSLKTDNRPEKSKCK
                PLWGKVFYLDLPSVTISEKLQKDIKDLGGRVEEFLSKDISYILSYLSNKKEAKFAQTLGRI
                SPVPSPESAYTAETRSPHPSDGSSFKSPDTVCLSRGKLLVEKRAIKDHDFIPSNILS
                NALSWGVKILHIDDIRYIEQKKKELYLLKKSSTSVRDGKRVGSGAQKTRTGRLKKP
                FVRVEDMSQLYRPFYLQLTNMPPIINYSIQKPCSPPDVKPSSMQKQTQVKLRIQTDGD
                KYGGTSIQLQQLKRRKKKGYCECLQKYEDLETHLLSEGHRRNFAQSNQYQVVDDIVSKL
                VEDFVEYEKDTPKKRKIKYSVGSLSPVSASVLKKTEQKEVELQHISQKDCQEDDTTV
                KEQNFLYKETQETEKKLLFISEPIPHPSNELRGLNEKRMSNKCCSMLSTABDDIRQNFTQ
                LPLHKNKQECILIDISEHLSENDLEELRVDHYRKCNIQASVHVSDFSTDNSGSGPKQKS
                DTVLFPAKDLKEKDLHSIFIHDSGLIITINSSQEHLTVQAKAPHTPPEPNECDFKNM
                DSLPSGKIHRKVKIILGRNRKENLEPNAEFDKTTEFIIQEENRICSSPVQSLLDLFQT
                SEEKSEFLGFTSTYEKSGICNVADIWEEENSDNLLTAFFSPSTSTFGF"
BASE COUNT      891 a    445 c    475 g    649 t
ORIGIN
```

```
Query Match     85.5%;  Score 2324;  DB 9;  Length 2460;
Best Local Similarity   97.5%;  Pred. No. 0;
Matches 2395;  Conservative  0;  Mismatches  0;  Indels 61;  Gaps 1:

Qy   319 cttcctcctccgcggccttgggagcggatccggccttgggagcggatccggccccgggaaaccgacctgcagacgcggt 378
         ||||||||||||||||||||||||||
Db     1 CTTCCTCCTCCGCGGCCTTGGGAGCCTTGGGAGCGGATCCGGCCCCGGGAAACCGACCTGCAGACGCGGT 60

Qy   379 acctctactgcgtagaggccggtagctggccggagaggagaggagaggagaggagaggacccggagaggcgcggt 438
         |||||||||||||||
Db    61 ACCTCTACTGCGTAGAGGCCGGTAGCTGGCCGGGAGAGGAGAGGACCCGGAGAGGCGCGGT 120

Qy   439 gccggggaagccgtgcttgcttccggggctgcgccggtgcgccgggtgcgacatg 498
         |||||||||||
Db   121 GCCGGGGAAGCCGTGCTTGCTTCCGGGGCTGCGCCGGTGCGACACTTTCTCCGGACCCAGCATG 180

Qy   499 taggtgccgggcgcgactcaagtccatgaacctccgcagagcgaaggaacatt 558
         |||||||||||
Db   181 TAGGTGCCGGGCGCGACTCAAGTCCATGAACCTCCGGAGATCCACAGTAAAGGACATTT 240

Qy   559 ccaggtggaatccaagtcaaaatgaaaaatagaaaaacaagaaccatctctgaaatctgaaaac 618
         |||||||||||
Db   241 CCAGGGTGGAATCCAAGTCAAAATGAAAAAAAACAGAACCATCTCTGAAATCTCTGAAAAAC 300

Qy   619 tgataacaggcccgaaaaatccaaaatgtaagccacacttgggaaggaaagtgtatttacctga 678
         |||||||||||
Db   301 TGATAACAGGCCCAGAAAAATCCAAAATGTAAGCCACTTTGGGAAGGAAAGTGTATTTTACCTTGA 360

Qy   679 cttaccttctgtcacccatctctgaaaaacttgaaaaagggacoattaaggagatctggagggcg 738
         |||||||||||
Db   361 CTTACCTTCTGTCACCATATTCTGAAAAAAACTTCAAAAGGGACGACTTAAGGAGATCTGGGAGGGCG 420

Qy   739 agttgaagaattctcagcaaagatatatacagttatctttattcttcaaatataggaagctaa 798
         |||||||||||
Db   421 AGTTGAAGAATTTCTCAGCAAAAGATATACAGTTATCTTTATTCTTCAAATAAGGAAGAAGCTAA 480

Qy   799 atttgcacaaacttgggtcgaattttctcctgtaccaagtccgatatgtcgcatataagctgc 858
         |||||||||||
Db   481 ATTTTGCACAAACCTTGGGTCGAATTTCTCCTGTACCAAGTCCGATATCTGCATATAACTGC 540

Qy   859 agaaaaccacttcacctcatccagcacgtcgatgatggaagttcactcaccaagacacagt 918
         ||||||||||||
Db   541 AGAAACCACTTCACCTCATCCAGCCAGCCATCAGTTCATTAAGTCACCAGAACAGT 600

Qy   919 gtgtttaagcagagagaaaattattagttcaaggactatcaagacctgaattttattcc 978
         ||||||||
Db   601 GTGTTTAAGCAGAGAGAAAATTATTAGTTGAAAAAGCTATCAAGGACCCATGATGATTTTATTCC 660
```

```
Qy   979 ttcaaatagttattattcaaatgcccttgtcatggggagggtaaaaatttcatattgatga 1038
         |||||||||||
Db   661 TTCAAATAGTTATTATTCAAATGCCTTGTCATGGGGAGGGTAAAAAGG 720

Qy  1039 cattagatactacattgaaacaagaagaagagaagagttgtatttactcaagaaatcaagtac 1098
         |||||||||||
Db   721 CATTAGATACTACATTGAAACAAGAAAAGAGGGGGCAAAAGAGAAAAGAACAGGAAG 780

Qy  1099 ttcaagtaagagagtggggcaagaagaagttggtagtggtgcacaaaagaaaggaag 1158
         |||||||||||
Db   781 TTCAGTAAGAGAGATGGGGGCAAAAGAGTTGGTAGTGGTGCACAAAAAGAACAGGAAG 840

Qy  1159 actcaaaaagccttttgtaaaggtggaagatgagcca--------- 1197
         |||||||||||
Db   841 ACTCAAAAAGCCTTTTGTAAAGGTGGAAGATGAGCCAAACTTTTATCT 900

Qy  1198 ---------aagccctgcagtccatttga 1217
              ||||||
Db   901 TCAGCTGACCAAATATGCCTTTTATATAAATTATTCTTATTCGAAGACCCCTGCAGTCCATTTGA 960

Qy  1218 tgtagacaagccatctagtatgcaaaagtcaaaactcaggttaaactaagaatccaaaacaga 1277
         |||||||||||
Db   961 TGTAGACAAGCCATCTAGTATGCAAAAGTCAAAACTCAGGTTAAACTAAGAATCCAAAACAGA 1020

Qy  1278 tggcgataagtatggtggaacctcaattcaaatccagttgaaagaagagaagaaaaagg 1337
         |||||||||||
Db  1021 TGGCGATAAGTATGGTGGAAACCTCAATTCAAATCCAGTTGAAAGAAGAGAAGAAAAAAGG 1080

Qy  1338 atattgtgaatgttgcttgcagaaaatatgaagatctagaaactcaacttctaagtgagca 1397
         |||||||||||
Db  1081 ATATTGTGAATGTTGCTTGCAGAAAATACTCAGAATCTCACCTTCTTAAGTGACCA 1140

Qy  1398 acacagaaaactttgcacagatgaaccagtataaagaatctaaagtgtttgatgatattttctga 1457
         |||||||||||
Db  1141 ACACAGAAAACTTTGCACAGAGTAACCAGTATCAAAGTGTTTGATGATATTGTATCTAAGTT 1200

Qy  1458 agtttgacttctgtggaatatgaaaggacacctaaaaagaatgaataataacg 1517
         |||||||||||
Db  1201 AGTTTTTGACTTTGCTGGAATATGAAAGGACACCTAAAAAGAGAATAAAATACAG 1260

Qy  1518 tgttgatccttctctcgtttctcctgttcctgcaagtgtcctgaaaaagactgaacaaaagaaa 1577
         |||||||||||
Db  1261 TGTTTGGATCCCTTTCTCGTTTCTGCAAGTGTCCTGAAAAAGACTGAACAAAAGGAAA 1320

Qy  1578 agtggaattgcaacatatttctccagaaagattgccaggaagcaagatgatacaacagtgaagga 1637
         |||||||||||
Db  1321 AGTGGAATTGCAACATATTCTCAGAAAGATTGCCAGGAAGCAAGATGATACAACAGTGAAGGA 1380

Qy  1638 gcagaatttcctgtatcaaagagaccccaggaaactgaaaaagctctgttatttcaga 1697
         |||||||||||
Db  1381 GCAGAATTTCCTGTATAAAGAGACCCAGGAAACTCAAGAAGGTCCTGTTTATTTTCAGA 1440

Qy  1698 gcccatcccccaccctcaaatgaattgagagggcttaatgagaaatgagtaataaatg 1757
         |||||||||||
Db  1441 GCCCATCCCCCACCCTCAAATGAATTGACAGGGCTTAATGAGAAAATGAGTAATAAATG 1500

Qy  1758 tccatgttacagctgaagtgacagcagaattttacacagctacctctcca 1817
         |||||||||||
Db  1501 TCCATGTTACAGCTGAAGTGACAGCAGAATTTTACACAGCTACCTCTCACA 1560

Qy  1818 taaaaaacaaacgaatgcaattcctgacatttccgacaacacattagtgaaaatgactt 1877
         |||||||||||
Db  1561 TAAAAAACAAACGAATGCAATTCCTGACATTTCCGACAACACATTAAGTGAAAATGACTT 1620

Qy  1878 agaagaactaaggggtagatcactatatatatgtaacacagcaatctgtacagtttctga 1937
         |||||||||||
Db  1621 AGAAGAACTAAGGGTAGATCACTATATATGTAACATACAGCAATCTGTACAGTTTTCGA 1680

Qy  1938 tttcagtacagatataatgtggatctcaaccaaacagaaagtcagaagatcagaaactgttgctttttcc 1997
         |||||||||||
Db  1681 TTTCAGTACAGATATAATGTGGATCTCAACCAAACAGAAAGTCAGATACTGTTGCTTTTCC 1740
```

```
Qy  1998  agcaaaggatctcaagagaaagaggacctccatcatcaattcattcatgttgttctggtctgat  2057
          |||||||| ||||||| |||| ||||||||| ||| ||| ||  || |||| |||||||||| |||
Db  1741  AGCAAAGGAGATCTCAAGCAAAAGGACCTTCATTCAATATTTACTCAGATTCTGGTCTCGAT      1800

Qy  2058  aacaataaacagttcacaagcaccctaactgttcaggtcaaaaggctccattccatactcc      2117
          |||||||| |||||||||||| ||||||||||||||| ||||| ||||||||||||||||||
Db  1801  AACAATAAAACAGTTCACAAGAGCACCTAACTGTTCAGGCAAAGGCCTCCATTCCATACTCC     1860

Qy  2118  tcctgaggaacccaatgatgtgacttcaagaatatggatagtgttacccttctcggtaaat     2177
          |||||||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  1861  TCCTGAGGAAACCCAATGATGTGACTTCAAGAATATGGATAGTTTACCTTCTCGGTAAAT      1920

Qy  2178  acatctgaaaagtgaaaataatatattaggacgaaaataagaaaaagagaaaatctcgaacca   2237
          ||||| |||||||| |||||||| ||||||||||||||| |||||||||||||||| ||||| |
Db  1921  ACATCGAAAAAGTCAAAATAATAATTATTAGGACGAAAATAGAAAAAGAGAAAATCTGAAACCAAAAGC  1980

Qy  2238  tgaattcgataaagaactgaatttattcacaagaagaaaacagaatttgtagttcacc       2297
          |||||| |||||| |||| |||| |||||||||||||||||||||||||||||||||||
Db  1981  TGAATTCGATAAAGAACTGAATTTATTACACAACAAGAAAACAGAATTTGTCAGTTCACC      2040

Qy  2298  ggtacagtcttactagacttgttctcagactagtgaagagaaatcagaatttttgggtttt    2357
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GGTACAGTCTTTACTAGACTTGTTGTTCTCAGAGTGAAGAGAAATCAGAATTTTTGGGTTT     2100

Qy  2358  cacaagcttacacagaagaagatcgggtatatgcaatcgttttagatatttgggaaggagaaa  2417
          ||||||||||||||||||||| ||||||| |||||||||||||||| ||| |||||| ||||||
Db  2101  CACAGCTTACACAGAAAAAGAGTGGTATATGCAATCGTTTTAGATATTATTGGGAAGAGGAAA    2160

Qy  2418  ttcagaataatctgttacacagcgttttctctgtccccctcaacttctaacattactggctt    2477
          ||||| |||| ||||||||||||| |||||||||||||| |||||||||||||||||||||||
Db  2161  TTCAGATAATCTGTTTAACAGCGGGTTTTCTCTGTCCCCTCAACTTCTACATTACTGGCTT      2220

Qy  2478  ttagaatttaaaaaatgcaatactttcagaagtgataaggatcatatatctttgaaatttt    2537
          |||| |||||||||||||||||||||||||||||||| |||||||| ||||||||||||||
Db  2221  TTAGAATTTAAAAAATGCATACTTTCAGAAGTGATAAGGATCATATATTCTTTGAAATTTTT    2280

Qy  2538  ataaatatgtatggaaattcttagatttttttaccagcttttgttttacagacccaaaatgt   2597
          ||||||||||| ||||||| |||||||||||||||||| |||| |||||||||||||||||
Db  2281  ATAAATATGTATGGAAATTCTTAGGATTTTTTTACCAGCTTTGTTTTACAGACCAAAATGT     2340

Qy  2598  aaaatattaaaaataaatatttgcaatctttctcacagaaattgaatacactgttaaagaaaat  2657
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||||| ||||||
Db  2341  AAATATTAAAAATAAATATTTGCAATCTTTCTCACAGGAATTGAATACCTGTTAAAGAAAAAT   2400

Qy  2658  tacagaataaactgtgactggtcttgtttttacattaaaaaaaaaaaaaaaaaa            2713
          ||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db  2401  TACAGCAATAAACTGTGACTGGTCTTGTTTTTACATTAAAAAAAAAAAAAAAAA            2456
```

RESULT 5

```
LOCUS       AC018977/c    236236 bp    DNA    HTG    26-JUL-2001
DEFINITION  Homo sapiens chromosome 10 clone RPl1-386C23, WORKING DRAFT
            SEQUENCE, 12 unordered pieces.
ACCESSION   AC018977
VERSION     AC018977.6  GI:15021987
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 236236)
AUTHORS     Smith,D.R.
TITLE       Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 236236)
AUTHORS     Smith,D.R.
TITLE       Direct Submission
JOURNAL     Submitted (25-DEC-1999) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Jul 26, 2001 this sequence version replaced gi:14091812.
```

```
            ----- Genome Center
                 Center: Genome Therapeutics Corporation
                 Center code: GTC
                 Web site: http://www.genomecorp.com/
                 Contact: gtc-seqcenter@genomecorp.com
            ----- Project Information
                 Center project name: hg092
            ----- Summary Statistics
                 Sequencing vector: N/A
                 Chemistry: Dye-terminator Big Dye; 100% of reads
                 Assembly program: Phrap; version 990315
                 Consensus quality: 228735 bases at least Q40
                 Consensus quality: 230209 bases at least Q30
                 Consensus quality: 231169 bases at least Q20
                 Insert size: 235235; sum-of-contigs
                 Quality coverage: 6.3x in Q20 bases; sum-of-contigs

            -----------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            -----------------------------------------------
            *        1     1146:   contig of 1146 bp in length
            *     1147     1246:   gap of unknown length
            *     1247     2466:   contig of 1220 bp in length
            *     2467     2566:   gap of unknown length
            *     2567     4025:   contig of 1459 bp in length
            *     4026     4125:   gap of unknown length
            *     4126     5165:   contig of 1040 bp in length
            *     5166     5265:   gap of unknown length
            *     5266     6309:   contig of 1044 bp in length
            *     6310     6409:   gap of unknown length
            *     6410     7552:   contig of 1143 bp in length
            *     7553     7652:   gap of unknown length
            *     7653     8716:   contig of 1064 bp in length
            *     8717     8816:   gap of unknown length
            *     8817    10307:   contig of 1491 bp in length
            *    10308    10407:   gap of unknown length
            *    10408    11913:   contig of 1506 bp in length
            *    11914    12013:   gap of unknown length
            *    12014    13363:   contig of 1350 bp in length
            *    13364    13463:   gap of unknown length
            *    13464   106707:   contig of 93244 bp in length
            *   106708   106807:   gap of unknown length
            *   106808   236236:   contig of 129429 bp in length
```

```
FEATURES             Location/Qualifiers
     source          1..236236
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-386C23"
                     /clone_lib="RPCI-11"
     misc_feature    1..1146
                     /note="assembly_name:Contig7"
     misc_feature    1247..2466
                     /note="assembly_name:Contig11"
     misc_feature    2567..4025
                     /note="assembly_name:Contig12"
     misc_feature    4126..5165
                     /note="assembly_name:Contig13"
     misc_feature    5266..6309
                     /note="assembly_name:Contig16"
     misc_feature    6410..7552
                     /note="assembly_name:Contig17"
     misc_feature    7653..8716
                     /note="assembly_name:Contig18"
     misc_feature    8817..10307
                     /note="assembly_name:Contig20"
     misc_feature    10408..11913
                     /note="assembly_name:Contig21"
```

```
misc_feature    12014..13363
                /note="assembly_name:Contig22"
misc_feature    13464..106707
                /note="assembly_name:Contig23"
                clone_end:T7"
misc_feature    106808..236236
                /note="assembly_name:Contig24"
                clone_end:SP6"

BASE COUNT    70014 a    44567 c    46458 g    74094 t    1103 others
ORIGIN

Query Match    73.1%;    Score 1986.6;    DB 2;    Length 236236;
Best Local Similarity    89.8%;    Pred. No. 0;
Matches 2268; Conservative    0;    Mismatches    69;    Indels 190;  Gaps  5;

Qy     254 gcggggccggccgggtgacggcggtttcaaatcttcaaccgccagccactgttgtgtt 313
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  218540 GCCGCGGCGGCCGGTGACGGCGGATTTCAAAACTTCAACCGCTGCCACTCGTTGTGATT 218481

Qy     314 tggccccttcctccggccctcggagaccggcgccccggaaaccgacctgcagac 373
           || |||||||||||||||||||| ||||||||||||||||||||||||| |||
Db  218480 TGGCCGCCTTCCTCCTCCTGGCCTTGGAGCCCGGGAAACCCGACCCACAGAC 218421

Qy     374 gcggtacctctactgcgtagagccgtagctgccggaagaggccgtcctgtc 433
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db  218420 GCCGTACCTCTACTGCGTAGAGGCCGTAGCTGGCGGAAGAGGCCGTCCTGTC 218361

Qy     434 aacaggcgcagggggaggccgtgctgcttcgcgggctgcccggctgcc 493
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db  218360 AACAGACGTGGGGAAGCCGTTGCTTCGTCGGGCTGCCCGGCTGCC 218301

Qy     494 gcatgtaggtgccgggcagctgccatgaactccggagccagtaaagga 553
           ||||||||||||||||||||||||||||||||| ||||||||||||||
Db  218300 GCATGTAGGTGTGCCCGGGCGGGACTGCCATGAACTCCGGAGCCAGTAAAGGA 218241

Qy     554 caattccaggggtcagaaagtcaaaaagtgaaaaaaaccatctctgaaatctctg 613
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  218240 CAATTCCAGGGTCAGAATCCAAGTCAAAAAAATGAAAAAACTCTCTGAAAATCTCTG 218181

Qy     614 aaaaactgattaacaggccagaaaatccaaatgtgaagccaagtccaaatttcac 673
           ||||||||||||||||||||||||||| |||||||||||| ||||||||||||
Db  218180 AAAACTGATTAACAGGCCAGAAAAATCCAAATATAAGCCAACTTGGGGAAAGTATTTTAC 218121

Qy     674 cttgactttacctttctgtcaccaatatctgaaaaacttcaaaaggaggatctggga 733
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  218120 CTTGACTTTACCTTCTGTCACCAATATCTGAAAAACTGAAAAGGATCTGGGA 218061

Qy     734 gggcggagtcagaagaatttctcagcaaagatatcagttatcttcaataagaaggaa 793
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  218060 GGGCCGAGTTCAGCAAAGATATCAGTTATCTTTATTTCAAATAAGAAGGAA 218001

Qy     794 gctaaaatttgcacaaaaccttgggtcgaattttctcctgtaccaagtctgcaat 853
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  218000 GCTAAATTTGCACAAACCTTGGGTCGAATTTTCTCCTGTACCAAGTCCAAAAT 217941

Qy     854 actgcagaaacacttcatccacgccatgatggaagttcattttcaccagac 913
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217940 ACTGCAGAAAACCACTTCATCCACGCCATGATGGAAGTTCATTTAAGTCACCAGAC 217881

Qy     914 acagtgtgtttagcgaggaggagaaaattattagttagtgaaaaagctatcaagagccatgatttt 973
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217880 ACAGTGTGTTTAAGCAGAGGAGAAAATTATATTAGTTGAAAAAAGCTATCAAGGACCATGATTTTT 217821

Qy     974 attcccttcaaatagtatatatatcaaatgcctgtcatgggggagtaaaatccttcatatt 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217820 ATTCCTTCAATATAGTATATATTCAAATGCCTTGTCATGGGGAGTAAAAATTCTTCATATT 217761

Qy    1034 gatgacaattagaatactacattgaacaaagaaaaagagagttgtattattccaagaaatca 1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217760 GATGACATTAGAATACTACATTGAACAAAGAGAAAAA 217703

Qy    1094 agtacttcagtaagaagagatggggcaaaagagagttggtagtgtgcacaaaaacaagaaca 1153
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

(right column)

```
Db  217702 TATCTTCAGCTGACCAAATATGCCTTTATAAATTATTCTATTCAGAAGAGCCCTGCAGTCCA 217643

Qy    1154 ggaagactcaaaaagccttctgtaaagtggaagatatgagcca 1197
           |||||||||||||||||||||||||||||||||| ----------
Db  217642 GGAAGACTCAAAAAGCCTTTTGTAAAAGTGGAAGATATGAGCCAACTTTATAGGCCANTT 217583

Qy    1198 ------------------------------------aagccctgcagtcca 1212
                                               |||||||||||||||
Db  217582 TATCTTCAGCTGACCAAATATGCCTTTTATAAATTATTCTATTCAGAAGAGCCCTGCAGTCCA 217523

Qy    1213 tttgattagacaagcatctagtatgcaaaagcaaactcaggttaaactagaaatccaa 1272
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217522 TTTGATGTTAGACAAGCCAATCTAGTATGCAAAAGCAAAACTCAGGTTAAACTAGAAATCCAA 217463

Qy    1273 acagaatggcgataagtaagtatggtggaacctcaattcaacctcagtgttgaagaagaagaaa 1332
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217462 ACAGATGGCGGATTAAAGTAGTGGAAATCAACTCAAGTTGAAGAAGAAGAAA 217403

Qy    1333 aaaggatatattgtgaatgttgcttgcaagaaatctgaaactcaccttcaagt 1392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217402 AAAGGATATTGTGAATGTTGCTTGCAGAAATATGAAGGATCTAGAAACT 217353

Qy    1393 gagcaacacagaaaacttgcacagagtaaccagtatcaaagtgtgttgatgatattgtatct 1452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217353 ------------------------------------------------------ 217353

Qy    1453 aagttagttttgactttgtgaatatgaaaaggacacactaaaaagaaagaataaaa 1512
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217352 AAGTTAGTTTTTGACTTTGTGGAATATGAAAAGGACACACTAAAAAGAAAGAATAAAA 217348

Qy    1513 tacagtcgttggatcccctttctcctgtcttctgcaagtgtcctgaaaaagactgaacaaag 1572
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217347 TACAGTGTTGGATCCCTTTCTCCTGTCTTCTGCAAGTGTCCTGAAAAAGACTGAACGAAAG 217288

Qy    1573 gaaaaagtggaattgcaacatattttctccaggaaagaagagagattcctcagaaaagaaaag 1632
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217287 GAAAAAGTGGAATTGCAACATATTTTCTCAGAAAGGAAGAGGATGATTATAACGGGTG 217228

Qy    1633 aaggagcagaagatttccctgtataaagagacccggaaactgaaaaaagctctgtttatt 1692
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217227 ATAGAGCAGAGAATTTCCCTGTATAAAGAGACCCAGGAAACTGAAAAAAA-CTTCTGTTTATT 217169

Qy    1693 tcagagcccatcccccacccttcaaatgaagtgagggcttaatgagagaatgagtaat 1752
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217168 TCAGAGCCCATCCCCCACCCTTCAAAAGTGAGACTGAGAGGGGCTTAACAGAGGAAATGAGTAAT 217109

Qy    1753 aaatgttccatgttaagtcgaagatgagagatgacagacaagctacct 1812
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db  217108 AAATGTTCCATGTTAAGTCGAAGATGACATAAGCTTGAACTACTAGCTACCT 217049

Qy    1813 ctacaataaaacaaacagggaatgcatcttgcatctccggaacacacattaagtgaaat 1872
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  217048 CTACATAAAACAAACAGGGAATGCATCTTGGACACACATTTTACACAGCTACCT 216989

Qy    1873 gacttagaagaactaagggtagatcacttaacagaaatgtaacatacagaagtcagaacatgtt 1932
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  216988 GACTTACAGAGAACTAAGGATAAGATAGATAGTAACTAACAGGCATGCTGTTTCANGTT 216929

Qy    1933 tctgattccagtacagatataagtggatctcaaccaacaaacagaagtcagatactgtgctt 1992
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  216928 TCTGATTTCTGTCACAGATAAATAGTGCATCTCAACCAAAAAACAGAAGTCAGATCCTGTGTGTT 216869

Qy    1993 tttcccgaaaggatctcaagagaaaaggaacctcattcaaataatatttcactctcatcctggt 2052
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  216868 TTTCCAGCAAAAGGATCTCAAGGAAAAAGACCTTCATTCAAATAATTTACTCGATTCTGGT 216809

Qy    2053 ctgataacaataacagttcacaaagaagagcaacctaaactgttcaggcaaaggcctccattcat 2112
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  216808 CTGATAACAATAAACAGTTCACAAAGACATCTTCACAAGAGCAACCTAAACTGTTCCAACAC 216749

Qy    2113 agtacttcagtaagaagagatggggcaaaagagagttggtagtgtgcacaaaaacaagaaca 2172
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 216748 ACTCCTCTGAGGAGACCCATGAATGTGACATCAAGAATATGGATAGTTTACTTCTGGT 216689
Qy   2173 aaaatacatcgaaaaggtgaaataatatattaggacgaagaagaaagaaatctggaacca 2232
          ||||||||||||||| ||||||  | ||||||||||| |||||| |||||||||||||||
Db 216688 AAAATACATCGAAAAGTGAAAATATTATTTAGGACAAAAATAGAAAAATCTGGAACCA 216629
Qy   2233 aatgctgaattgataaaaagaaagaaactgaatttattaccacaagaagaaaacagaga 2292
          |||||||||| |||||||| || ||||||||| ||||||||| |||||| ||||||||
Db 216628 AATGCTGAATTTGATAAAAGAAGAAACTGAAATTTATTACAACAAGAGAAAACAGAGA 216569
Qy   2293 tcaccggtacagtcttactagaactgttctcagacagtgaagaagaaaatcagaatttg 2352
          |||||||||||||||| || |||| |||||||||||||||||| |||||||||||||
Db 216568 TCACCGGTACAGTCTTTACCAGACTTGTTCTCAGACAGTGAAGAGAGAAAATCAGAATT 216509
Qy   2353 ggttcacaagtcacacagagaaaggagtgtgtatatgtatgcaatgtttttgatattg 2412
          |||| |||| ||| ||||| | || |||| ||| || ||||||| |||||| |||| |
Db 216508 GGTTTCACAACGTATACAGAAGAGAGTGGGTATATGCAATGTTTTTAGATATTTTGGGAAGAG 216449
Qy   2413 gaaaattcagataatctgttaacacgcgtgtttttcctcgtcccccttcaacttctacactt 2472
          ||||||||||||||||||||||||| ||||||||| ||||| || ||||||||||||  ||
Db 216448 GAAAATTCAGATAATCTGTTAACACGTGTTTTTCTCGTCCCACTTCAACTTCTACATTTACT 216389
Qy   2473 ggctttagaattaaaaatgcatacttttcagaagtgataagtgtcatattcttgaaa 2532
          ||||| ||||| |||||| |||||||||||||||||||| | |||||||||||||||
Db 216388 GGCTTTTAGAATTTAAAAAATGCATACTTTTCAGAAGTGATAAGGATCATATTCTTGAAA 216329
Qy   2533 tttttatataatgtatgggaaatctcttagga-ttttttttttaccagcttttgttta 2591
          |||||||||||| ||||||||||  | |||||  ||||||||||||||||||||||||
Db 216328 TTTTTATATAATAATGTATGGGAAATTCTTAGGATTTTTTTTTACCAGCTTTTGTTTTACAGACCC 216269
Qy   2592 aaatgtaaataa-tataaaaataaaattgaattttgaaatgaatacctgttaaag 2651
          ||||||||||||  || ||||| |||||||||||||||||| |||||||  |||||
Db 216268 AAATGTAAATAATAAAATATTTAAAAATAAAATTGAATTTTGCAATTTCTTACAGGA 216209
Qy   2652 aaaaattcacagaataaaactgtgactggtcttgtgttcttacattacaagaattg 2711
          ||||| ||||| || ||||| |||||||||| ||||||||| |||  || ||||||
Db 216208 AAAAATTACAGAATAAAACTTGTGACTGGTCTTGTTAATTTTTACAGTTTAAAAAGAAGAATTT 216149
Qy   2712 aacccga 2718
          ||| ||
Db 216148 GACACAA 216142

RESULT  6

LOCUS       AF292400    2458 bp    mRNA            ROD       21-MAR-2001
DEFINITION  Cricetulus griseus DBF4/ASK mRNA, complete cds.
ACCESSION   AF292400
VERSION     AF292400.1  GI:13398340
KEYWORDS    .
SOURCE      Chinese hamster.
  ORGANISM  Cricetulus griseus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 2458)
  AUTHORS   Guo,B. and Lee,H.
  TITLE     Cloning and characterization of Chinese hamster homologue of yeast
            DBF4 (ChDBF4)
  JOURNAL   Gene 264 (2), 249-256 (2001)
  MEDLINE   21147935
REFERENCE   2  (bases 1 to 2458)
  AUTHORS   Lee,H. and Guo,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2000) Research, Northeastern Ontario Regional
            Cancer Centre, 41 Ramsey Lake Road, Sudbury, ON P3E 5J1, Canada
FEATURES             Location/Qualifiers
     source          1..2458
                     /organism="Cricetulus griseus"
                     /db_xref="taxon:10029"
     CDS             195..2225
                     /codon_start=1
                     /product="DBF4/ASK"
                     /protein_id="AAK21956.1"
                     /db_xref="GI:13398341"
                     /translation="MNSGAMRIHSKGHFQQGGIGVKNEKNRPSLKSLKTDNRPEKSKCK
                     PLMEKVFYLDLPSVTISEKLQKDIRKDLGQGRVEEFLSKDISYLVSNNKEAKYAQTLGQV
                     SPVPSPESAYTAETTSPHPSHDGSSFKKSPDRVCLSRQKLLAEKAVKDHDFIPANSLLS
                     NALTWGVKILHIDDIRYYIEQRKKELCSLKKSSTSYRDSKRKATTIIQKARTGRLKFP
                     FLKVEDVNRSYRPFYLQLTSVPSINYATHKPCSPFDIEKPSSVQKKQAQPKPRNTDGD
                     KCGGTPVQLQLREKRKRGYCECCLLQKYEDLETHLLSEKHKNFAQSNQYQVDDIVSKL
                     VFDFEVEYERDTPFKKIRKYSVGSFSSSVSANVLKNTEPKEKLQLPFIPQKDMVESNGQL
                     PEEIFQCEDIQCEDIQKPEQRLVILASEPMSYSSTGLKGRDEKAASMLNASEPDIKQFF
                     TQLPPCKNEQGILYSEHRLIIINRNDLEQRVGDSVGVPRSCVQVSHLSPENSLPQPK
                     LTADTTHFSAKDLQEKDLHPVFGHDSDLVTLNTSKEQLTYKAGTPSCGPQQPNECDTE
                     NTDNLPCCKIQRKVRLLLGQKKRNVDPSAELDKKRTEFLPMCEDRTCGGSPVQSLLDLF
                     QTSGEKSDFLGFTSTTENSGLCDVLDWEDENSSSLLLSTPFSSPSASTPIGF"

BASE COUNT      807 a    467 c    533 g    651 t
ORIGIN

Query Match  52.5%;  Score 1427.4;  DB 10;  Length 2458;
Best Local Similarity  77.5%;  Pred. No. 1.1e-260;
Matches 1916;  Conservative 0;  Mismatches 441;  Indels 114;  Gaps 11;

Qy  330 gcgcccttggagccgggatccggcccgcccccggaaaccgacctgcagaagccgcggggggaag 389
        ||||||| |||||||||||||| ||||||||||||||||||||||||||| |||||||||||||
Db    8 GCGCCCTTGGACCCGGATCCGGCCCCGGCAAAACCCGAGCTGCAGAGCGCGGTACCTCTACTGC 67

Qy  390 gtagagcggccatagctcggcgcgagaaggagaaggcgggccgtccctgtcaacaggccgggggaag 449
        ||||||| ||||||||||| |||||||||||||||||||||||||| ||||||  |||||||||
Db   68 GTPAGAGCCCGGTAGCTGGGCGGAAGGAGAGGCGGGCGGTCCGTGTCAACAGG-CGGGGGAAG 126

Qy  450 ccgtgcttcgcggcctggcgccgggtgcgacaacttctcccggaccagcatgtaggtgccgggg 509
        |||| ||||| |||||||||||||||||||||||||| |||||||||||| |||||||||| ||
Db  127 CCGTGCTTCGCGCGGCTGGCGCCGGGTGCGACAACTTCTCCGGACCAGCATGTAGGTGCCGGG 186

Qy  510 cgactgccatgaacctccggagccatgaggatccacagtaaaggggacattccaggtggaa 569
        ||||||||| |||||||||||||||||||||||||| || |||||||| ||||||||||||
Db  187 CGACTGCCATGAACCTCCGGAGCCATGAGGATCCACAGTAAAGGACATTCCAGGTGGAA 246

Qy  570 tccaagtcaaaatgaaaaaaacagaacccatctctgaaatctctgaaaactgataacaggc 629
        |||||||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db  247 TCCAAGTCAAAAATGAAAAAAACAGAACCCATCTCTGAAAATCTCTGAAAAACTGATAACAGGC 306

Qy  630 cagaaaaatccaaatgtaagccaacttgggaaaagtgtattttacctgactacctctg 689
        ||||||||||||||||||||||||||| |||||||||||||||| ||||||||||||||
Db  307 CAGAAAAATCCAAATGTAAGCCAACTTTGGGAAAAGTATATTTTACCTTGACTTACCTTCTG 366

Qy  690 tcacccatactgaaaaactcaaaaggacattaagggagtcgggaggcgggaggttgaagaat 749
        |||||||||||||||||||||||||||| |||||||||||| || ||||||| ||||||||
Db  367 TCACCATTACTGAAAAAACTTCAAAAGGACATTAAGGGAGTCGGGAGGCGGGAGGTTGAAGAAT 426

Qy  750 ttcttcagcaaagatatcagttatcttattccaaaataagaagaggaagctaaaatttgacaaa 809
        ||||| |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  427 TTCTCAGCAAAGATATCAGTTATCTGTTCAAAATAAGAAGAGGAAGCTAAAATATGCACAA 486

Qy  810 cctggtcgaattctcctgtaccaagtccagaatctgcatatactgcagaaaccactt 869
        |||||||| |||||| ||||||||||||||||||||||| ||||||||||| ||||||
Db  487 CACTGGGTCAAGTTTCTCCTGTACCAAGTCTGCATATACTGCAGAAACCACCT 546

Qy  870 cacctcaatcccagccatgatgaagttcaattaagtcacccagacacaagtgttgttaagca 929
        |||||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||
Db  547 CACCTCATCCCAGCCATGATGGAAGTTCCTTAAGTCACCCAGACACAGTGTGTTTGAGCC 606

Qy  930 gaggaaaaatatttagttcaaaaagctatcaaggagccaacccagattttattccctcaaaatagta 989
        | || |||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||
Db  607 GAGGGAAGTTATTAGCTGAAAAAGCTGTCAAGGATCATGATTTTATTCCTGCTAATAGTA 666

Qy  990 tattatcaaatgcctcatgggggtaaaaatccttcatatccattggtgacattagtatact 1049
        ||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db  667 TATTATCAAATGCCTCACATGGGACTAAAGACTCTTCATATATTGATACACATTAGTAGTACT 726

Qy 1050 acattgaacaaaagaaaaaaaagagttgtatttactccaagaaatacattcaagtaagag 1109
        ||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||
Db  727 ACATTGAACAGAGAAGAAGAAAGAAGAGTTGTCACTCAAGAAATACAAGTACTTCGTGTGAGGG 786
```

```
Qy  1110 atgggggcaaaagagttggtagtggtgcacaaaacaagaacacggaagaactcaaaagc 1169
         |||   ||||||| || |  |||  | ||  ||| ||||| ||| || || ||||||||
Db   787 ATTCGGGAAAAAGCTGGCACTACAATACAAAGGCAGGGAGGACTCAAAAAGC          846

Qy  1170 ctttgtaaaggtggagtataga---------                               1193
         ||||  ||||||| |  ||||
Db   847 CTTTTTTAAAGGTTGAAGAGTGAACCGGTGTCTTACCTTCAGGCCATTCTTACCTTCACAGGCCA 906

Qy  1194 ------gccaaagccctgcagtccattgatgtagacaagc                       1228
               ||||||  | ||||||||| || |||||| || |
Db   907 GTGTGCCTTCTATAAAACTATGCTACGCACAAGGGCTTGCAGCCCATTTGCACTAGAGAAGC 966

Qy  1229 catcctagtatgcaagagcaaaactccaggttaaactaacagacagatggcgataagt     1288
         ||||   |||| | |  ||||||   |||||| ||||||| ||||| ||  |||||
Db   967 CATCCACACCGTCCAAAAACAGCTTCAGCCGTCTAAACCCAAGCTCAGGGGAAATCAGTT    1026

Qy  1289 atggtggaacctcaattcaactccagttgaaagaaagaaagagatattgtgaat        1348
         |||||||||| |||| ||||||||||||||||||||| ||||||||||   |||
Db  1027 GTGGTGGAACCCCAGTTCAACTCCAGTCAACAATGAAAGAAAAAGGATATTGTGAAT      1086

Qy  1349 gttgcttgcagaaatatgaagatctcagagaaactcacccttctacagagtgagcaaact 1408
         |||||||||||||||||||||| |||| |||||  |||   | || || ||| |||| ||
Db  1087 GTTGCTTACAGAAATATCAAGAGCTCAGAGAAATTCACTGGAAGAGAAAAAGGATAATACT 1146

Qy  1409 ttgtcacagaagtaaccagtatcaaagtygttgtgatattgtatctaagttagtttttgact 1468
         ||| ||||| | |||| || ||| ||||| || || |||||||||| |||| |||  |||
Db  1147 TCGGCACAGAGTAAACCAGTAATCAGAAGTGTTGATGATATTCACCTTTTGAGTGAGAAACT  1206

Qy  1469 ttgtggaatatgaaaaggcacacagaaaaagaaaataacacagtgtggatcc          1528
         |||||||| ||||| ||| ||  || ||||||| |||| ||| ||| |||||
Db  1207 TTGTGGAATATGAGAGAGCACAACAGAGAAAGAATATAAAACACGGTGGGTCTCCT        1266

Qy  1529 ttctcctgttctgtccaagtgtcctgaaaagactgaaacaaagtggaattgc          1588
         ||||| ||||| |||||| | ||||| | ||| ||||  ||||| ||||| ||
Db  1267 TTTCTTGTTCTGCAAATGAAGCACAACACTGAAACAAAGGATTAGTTTTGGCTCTT       1326

Qy  1589 aacatatttctcagaaagattgccaggaagatgaacaacagtgaaggaagcaga------- 1642
         |||| |||  |||| |||| | |||||||| | ||  ||| | | |||||| || 
Db  1327 AGCCCATTTCTCAGAAGGACAGGGTAGAAAGTAATGAAGAACTCCTGAAGGAGTAGAATTTTTC 1386

Qy  1643 ------atttcctgtataagagaccagaaactgaaaaaagctccctgttattt          1693
               ||||||| |||||||||||||||| |||| |||||||   ||| ||
Db  1387 AGTGTGAAGATATCCAGTGTGAAGATATCCGAACACAAAGGCTTGTGTTGGCTT        1446

Qy  1694 cagagcccatcccccacccttcaaatgaattgagagggcttaatgagaaatgagtaata   1753
         ||||  || ||     |||| ||| ||||  |||||| |||| |||||||| ||||
Db  1447 CAGAACCATGTCCTACTCTTCAACGGGATTTGAAAAGGACGGTGAGACAAGAAA--------G 1497

Qy  1754 aatgttccaatgttaagtacagtgaagatgacataagacaagcagaagctactctc      1813
         ||||||||||||||||||| || ||| | | ||| |  | || |||| ||   ||
Db  1498 CAGCCACCATGTTAAATGCCATGTAAACCTGACATAAAACAGAAGTTAACACAGCTACCTC    1557

Qy  1814 tacataaaaacaacacaggaatgcatctcttgacattccgaacacacattaagtgaaaatg 1873
         ||| || ||||||  || || ||| |||| ||  ||| | | ||| ||||| |||| ||
Db  1558 CATGTAAAAATGAACAGGAAGAAGTTCTTGAAGCATTAATTAATTAACAAGTGAAATG     1617

Qy  1874 acttaga--agaactaagggtagaatcactatataaatgtaacatacagccatctgtacatg 1930
         ||| |||  |||  |||  | ||  |  || |||||| | || |||| || |||| || 
Db  1618 GAAATGATTTAGAACAGAGGGTTGGTGACTCGTGAGGTGTCGCCACGGTCGTGTCATGTGG    1677

Qy  1931 ttcctgattcagataacagtaatatggatctcaaacaacagaagtcagatactgtgc     1990
         |||| |||||| | |||| || | ||| || |||||||||| |||||| || ||||
Db  1678 TTTCTCATCTCAGTCCAGTACAGACTGCTACCCCAACCAAAATTAACAGCATCATCAC     1737

Qy  1991 ttttttccagcaaggagatctcaaggaaaagacccttcattcatttactctgattctg    2050
         |||||||| |  ||||||   ||| |||||  || ||||||||  |  ||| ||| ||
Db  1738 ATTTTTCAGCAAAAGATTTTACAGAAAAAAGGACCTTCATTTGTTTGTCTCATGGATTCTG  1797

Qy  2051 gtctgataacaatataacagttcacaagacgcctaactgttcagcgaaagctccaattcc 2110
         ||| ||| |||||  ||||| | || ||||   ||| |||| || |||| |||||||||
Db  1798 ATCGGGTGACATTAAATACATCAAAGAAGAACAGTAAGCAGTCAAGGCCGGAACGCGCCATCCT 1857
```

```
Qy  2111 atactcctcctgaggaacccaatgaatgtgacttcaagaattatgatagtttacgttctg 2170
         || |||||| ||||||| ||| || |||| |||| |||| || || | |||  || |||
Db  1858 GTGGGTCCTC---AGCAGCCCAATGAGTGTGACACCGGAGAATCACAGATAACTTGCCTTGTG 1914

Qy  2171 gtaaaatacatcgaaaagtgaaaataaatattaggacgaaatagaaatcttggaac     2230
         || ||||||| ||| ||||| |||| ||| || | ||  |||||||| |||||||
Db  1915 GTAAGAATCCACGCAGGGAAAGTGAGACACTGTTAATATTAGGCCAGAG---AAAAAATGTGGACC 1971

Qy  2231 caaatgctgaatttgat---aaaagaactgaatttattaccacagaagaagaaacagattt 2287
         |||| ||||||||| ||   |||| |||||| ||| ||||| |  ||| ||| | |||||
Db  1972 CAAGTGCTGAATTTGATAAAAAGAGAACTGAGTTTTCCTATGTGTGTGGAAGACAGAGAACTT 2031

Qy  2288 gtagttcaccggtacagtgtcttttactcagtcgttccagactagtgaagagaaatcagaat 2347
         ||||  |||||| ||||||| ||||||| ||| || |||||| || ||||||||| ||||
Db  2032 GTAAATCCACGTGAAATTGAAAATAAATATTAGGACGAAATAGAAATCTTGGAC        2091

Qy  2348 ttttgggtttcacaagctcaacacagaaagaaagaggttagtatgtcaaagtgtttagatattt 2407
         |||||||| ||| |||||||| ||||| |||| ||| | ||  | ||| ||| ||||||
Db  2092 TTTTGGGTTTCACCAGCTACGCTACACTGGAGTCTCCTATGTGTGGAAAGGCTATGTGTTTGGG 2151

Qy  2408 aagaggaaaaatcagataatctgttaacagcgtttttctcgtcccccttcaacttctcat 2467
         |||  ||| | ||| ||||| | || || | ||||||||| | |||||||||| ||||
Db  2152 AGGATGGAAAATCAAGAAAATCCTGCTCTCAACACATTTTTCTCTCCCCTTCAGCTTCAT 2211

Qy  2468 ttactggcttttagaatttaaaaaatgcatacacttcagaagtgatataagtgtcatatct 2527
         || || ||||| |||||| |||||| |  || || |||||| | ||||| | ||| |||
Db  2212 TCATTGGATTTTAGTAGTATAAGTGTTTTCAGAAGTGTGGACAGTTTTAATTTTATTCTGAAT 2271

Qy  2528 tga---------aattttatataaaatcatgaaatcttaggagtcttaatataattctacaggaat 2575
         |||         ||||||| ||||||||  ||||||||| | || ||   ||| |  ||||
Db  2272 TATAAATACATATAATTCTTAGTAGTATGTTTTCAGAAGTGTGGACAGTTTTAATTTTATTCCAG 2331

Qy  2576 ctttgtttacagaccccaaatgtaaaatattttaaaatataactggaatttctacacagaat 2635
         ||||| |||||  |||| |||||||||| |||||||| ||||||||| ||||||| |||||
Db  2332 CTTTTGTTGCTAGAACCATATGTGAAAACCATATGTGAAAATAAATACTAAGA--TGAATGTTTATT 2389

Qy  2636 tgaatacctgttaaagaaaatttacagaatataaacttgtgactggtcttgttttacattaa 2695
         |||||||| || |||| |||||||| || |||| |||||||||||| |||||||  |||||
Db  2390 TGAGTAGCTGTTT--ACAAATTATACAATACAATTTTATTTATTAATCTTTGTGATTTTTATTTATTA 2447

Qy  2696 aaaaaaaaaa                                                     2706
         ||||||||||
Db  2448 AAAAAAAAAA                                                     2458
```

RESULT 7
LOCUS   AC005164/c
AC005164 123331 bp   DNA   PRI   03-FEB-2000
DEFINITION Homo sapiens BAC clone CTB-135C18 from 7q21, complete sequence.
ACCESSION   AC005164
VERSION   AC005164.1 GI:3242749
KEYWORDS   HTG.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 123331)
  AUTHORS   Kellen,J. and Burkhart,J.
  TITLE   The sequence of Homo sapiens BAC clone CTB-135C18
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 123331)
  AUTHORS   Waterston,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-JUN-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   3 (bases 1 to 123331)
  AUTHORS   Waterston,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (03-FEB-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT   -------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC

Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
------- Summary Statistics
Center project name: H_RG135C18

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping project (Eric D. Green, Director; John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see http://www.nhgri.nih.gov/DIR/GTB/CHR7, send mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
Clone CTB-135C18 is from the first release of the human BAC library CITB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992): U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (http://www.resgen.com).
VECTOR: pBeloBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP5-1189K17; the clone sequenced to the right is CTB-60N22, 200 bp overlap. Actual start of this clone is at base position 1 of CTB-135C18; actual end is at 123138 of CTB-135C18.

FEATURES   Location/Qualifiers
source   1..123331
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q21"
                     /clone="CTB-135C18"
                     /clone_lib="CITB-978SK-B"
repeat_region   36..78
repeat_region   235..322
                     /rpt_family="MER2_type"
repeat_region   1466..1636
                     /rpt_family="Alu"
repeat_region   1995..2122
                     /rpt_family="MIR"
repeat_region   2782..3080
                     /rpt_family="Alu"
repeat_region   3035..3134
                     /rpt_family="MaLR"
repeat_region   3136..3664
                     /rpt_family="MaLR"
repeat_region   3786..3855
                     /rpt_family="Retroviral"
repeat_region   3891..4128
                     /rpt_family="MaLR"
repeat_region   4476..4714
                     /rpt_family="MaLR"
repeat_region   4882..5225
                     /rpt_family="MIR"
repeat_region   /rpt_family="MaLR"

repeat_region   5762..5939
                     /rpt_family="MER1_type"
repeat_region   5950..6035
                     /rpt_family="Alu"
repeat_region   6350..6642
                     /rpt_family="MER1_type"
repeat_region   7010..7308
                     /rpt_family="Alu"
repeat_region   9806..10112
                     /rpt_family="L1"
repeat_region   10303..10498
                     /rpt_family="Alu"
repeat_region   10507..10807
                     /rpt_family="Alu"
repeat_region   10861..10978
                     /rpt_family="Alu"
repeat_region   11533..11836
                     /rpt_family="L1"
repeat_region   11959..12079
                     /rpt_family="Alu"
repeat_region   13392..14290
                     /rpt_family="Alu"
repeat_region   15571..16119
                     /rpt_family="Alu"
repeat_region   16142..16438
                     /rpt_family="Alu"
repeat_region   16500..16779
                     /rpt_family="L1"
repeat_region   16933..17077
                     /rpt_family="Alu"
repeat_region   17850..18392
                     /rpt_family="L2"
repeat_region   18460..18769
                     /rpt_family="L1"
repeat_region   18908..19366
                     /rpt_family="MaLR"
repeat_region   21782..21896
                     /rpt_family="Alu"
repeat_region   22048..22156
                     /rpt_family="L2"
repeat_region   24873..25250
                     /rpt_family="MER1_type"
repeat_region   26032..26309
                     /rpt_family="MaLR"
repeat_region   26342..26817
                     /rpt_family="L1"
repeat_region   27147..27355
                     /rpt_family="Alu"
repeat_region   27381..29443
                     /rpt_family="L2"
repeat_region   29444..29561
                     /rpt_family="MER1_type"
repeat_region   31770..31909
                     /rpt_family="MaLR"
repeat_region   32332..32720
                     /rpt_family="Alu"
repeat_region   32731..32953
                     /rpt_family="MaLR"
repeat_region   33094..33227
                     /rpt_family="L2"
repeat_region   33643..33941
                     /rpt_family="L1"
repeat_region   34000..34088
                     /rpt_family="Alu"
repeat_region   34233..34602
                     /rpt_family="MIR"
repeat_region   38520..38721
                     /rpt_family="MaLR"
repeat_region   38730..38849
                     /rpt_family="Alu"
repeat_region   40214..40605
                     /rpt_family="MER1_type"
repeat_region   41565..41638
                     /rpt_family="Alu"
repeat_region   /rpt_family="Alu"
repeat_region   /rpt_family="MaLR"
repeat_region   /rpt_family="MIR"
repeat_region   /rpt_family="Alu"
repeat_region   /rpt_family="L2"

```
repeat_region    /rpt_family="L1"     42099..42347
repeat_region    /rpt_family="MIR"    42470..42784
repeat_region    /rpt_family="L1"     42787..43087
repeat_region    /rpt_family="Alu"    43588..43892
repeat_region    /rpt_family="Alu"    44036..44116
repeat_region    /rpt_family="Alu"    44562..45028
repeat_region    /rpt_family="L1"     45191..51321
repeat_region    /rpt_family="L1"     52546..52617
repeat_region    /rpt_family="L2"     53330..53581
repeat_region    /rpt_family="L2"     54254..54761
repeat_region    /rpt_family="L1"     54759..55585
repeat_region    /rpt_family="L1"     55659..55957
repeat_region    /rpt_family="Alu"    58134..58298
repeat_region    /rpt_family="MIR"    59008..59181
```

```
Query Match     43.9%;  Best Local Similarity  99.7%;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;
Matches 1196;   Score 1194.2;  DB  9;  Length 123331;  Pred No. 2e-216;
```

```
Qy   1502 aaagaataaaatacagtgtttggatccccttgtttctgttcctgcaagtgtcctgaaaaga 1561
          |||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||
Db  94286 AAAGAATAAAATACAGTGTTTGGATCCCCTTGTTTCTGTTCCTGCAAGTGTCCTGAAAAGA 94227

Qy   1562 ctgaacaaaaggaaaaagtggaattgcaacaatatttctcagaaaagattgccaggaagatg 1621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94226 CTGAACAAAAGGAAAAAGTGGAATTGCAACAATATTTCTCAGAAAAGATTGCCAGGAAGATG 94167

Qy   1622 ataacaacagtgaaggagcagaattcctgtataaagaagaagaaacctgaaaaaagc 1681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94166 ATACAACAGTGAAGGAGCAGAATTTCCTGTATAAAGAGAAGAAACTTGAAAAAAGC 94107

Qy   1682 tcctgtttattcagagccccatcccccacccctcaaatgaattgagagggcttaatgaga 1741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94106 TCCTGTTTATTCAGAGCCCCATCCCCCACCCCTCAAATGAATTGAGAGGGCTTAATGAGA 94047

Qy   1742 aaatgaagtaagtaaaaagtgttccaatgttaagtacacagagtgacataagacagaatttta 1801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94046 AAATGAAGTAATAAAAAGTGTTCCAATGTTAAGTACACAGACTGACATAAGACAGAATTTTA 93987

Qy   1802 cacagctacctacataaaaacaaacagcatgttccgaaacacacat 1861
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  93986 CACAGCTACCTCACATAAAAACAAACAGCATGTTCCGAAACACACAT 93927

Qy   1862 taagtgaaatgactagaagaactaaggtagatcacatatacatacaggcat 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93926 TAAGTGAAATGACTAGAAGAACTAAGGTAGATCACATATACATACAGGCAT 93867

Qy   1922 ctgtacaagttctgattccagtacgacaagatataatagtggatctccaaccaaaacaagtcag 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93866 CTGTACAAGTTTCTGATTCCAGTACGACAAGATATAATAGTGGATCTCCAACCAAAACAAGTCAG 93807

Qy   1982 atactgtgtcttttccaagcaaaggatctcaaggagaaaggagacttcattcattcaataitactc 2041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93806 ATACTGTGCTTTTTCCAGAAGGATCTCAAGGAGAAAAGAGATCTCATTCAATATTTACTC 93747

Qy   2042 atgattcctgataacaacaaacagttcacaagagagcacctaactgttcaggcaaagg 2101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93746 ATGATTCTGTGTCTGATAACAATAAACAGTTCACAAGAGCACCTACTGTTCAGGCAAAAGG 93687
```

```
Qy   2102 ctccatccactccctcctgaggaacccaatgaatgtgactgtagaaatatgtagtt 2161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93686 CTCCATTCCATCCCTCCTGAGGAACCCAATGAACCCAATGAATATGGATATAGTT 93627

Qy   2162 tacctcctgtaaatacatcgaaagtgaaaatataattaggacgaaatgaaaagaa 2221
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93626 TACCTTCTGTAAATACATCGAAAAGTGAAAATAATATTAGGACGAAATACAAAAGAAA 93567

Qy   2222 atctggaacccaaatgctgaatttgataaaagatcgaattattaccacaagaagaaaaaca 2281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93566 ATCTGGAACCCAAAAGTGAATTTGATAAAAGATCGAATTTATTACACAAGAAGAAAAACA 93507

Qy   2282 gaatttgtagttcaccggttacagtctcttactcagacttgttccaggctagtgaagagaat 2341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93506 GAATTTGTAGTTCACCGGTTACACTTTTACTCAGACTTGTTCCAGGCTAGTGAAGAGAAAT 93447

Qy   2342 cagaatttttgggtttccacagctacacacagaaaaaagtagtatatgcaatgttttagata 2401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93446 CAGAATTTTTGGGTTTCACAGCTACAGTCTTTTACTCAGACTTGTTGTTGTTTCAATGTTTTAGATA 93387

Qy   2402 tttgggaaggagaaatcagagataactgttaacagcggttttctctcgtcccctccaactt 2461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93386 TTTGGGAAGGAGAAATTCAGAGATAAATCTGTTAACAGCGGTTTTCTCTGTCCCCTTCAACTT 93327

Qy   2462 ctacatttactggctcttagaatttaaaaaatgcataactctttcagaagaaagtgataaggatca 2521
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93326 CTACATTTACTGGCTTTTAGAATTTAAAAAATGCATAACTTGTCGTCGTTTGTTTTTACATTAATATA 93267

Qy   2522 tattcttgaaatttatatagtgatggaaatttcttaggatttttttttacagcttgt 2581
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93266 TATTCTTGAAATTTTTATATAGTGTATGGAAATTCTTAGGATTTTTTTACCAGCTTTTGT 93207

Qy   2582 ttacagacccaaatgtaaatattataaaaataaatatttgcaatttctacagaattgaata 2641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93206 TTACAGACCCAAATGTAAATATTATAAAAATAAATATTTGCAATTTCTACAGAATTGAATA 93147

Qy   2642 cctgttaaagaaaattacagaataaaactgttgactggtcttgttgttttacattaaaaaaa 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  93146 CCTGTTAAAGAAAATTACAGAATAAAACTTGTCGACTGGTCTGTTGTTTTTTACATTAATATA 93088
```

```
RESULT   8
AC024947
LOCUS       AC024947     163031 bp    DNA      PRI      24-AUG-2000
DEFINITION  Homo sapiens chromosome 01 clone RP11-200I10, complete sequence.
ACCESSION   AC024947
VERSION     AC024947.4  GI:9887599
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163031)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163031)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
REFERENCE   3  (bases 1 to 163031)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-AUG-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Aug 24, 2000 this sequence version replaced gi:8569066.
FEATURES             Location/Qualifiers
     source          1..163031
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="01"
```

Left column

```
Query Match      43.9%;   Score 1194.2;   DB 9;   Length 163031;
Best Local Similarity   99.7%;   Pred. No. 2.1e-216;
Matches 1196;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

BASE COUNT     48079 a   30305 c   32059 g   52588 t
ORIGIN

                    /clone="RP11-200110"
                    /clone_lib="RPCI-11"

Qy   1502 aaagaataaatacagtgttggatccttctcctgtttctgcaagtgtcctgaaaaga 1561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35370 aaagaataaatacagtgttggatccttctcctgtttctgcaagtgtcctgaaaaga 35429

Qy   1562 ctgaacaaaagaggaaaagtggaattgcaacaatatttctcctgttgccaggagatg 1621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35430 ctgaacaaaagaggaaaagtggaattgcaacaatattttctcagaaagattcc 35489

Qy   1622 atacacagtgaggagcagaagaattcctgtataaagagacccagagaaactgaaaaagc 1681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35490 atacaacagtgaagaggaggcagaatttcctgtataaagagacccaggaaaactgaaaaaaac 35549

Qy   1682 tcctgtttattcagagccgcccatcccccaccccttcaaatgaattgaagggcttaatgaga 1741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35550 tccctcgttttaattcagagcccatcccccacccccttcaaaatgaaatgaagggcttaatgaca 35609

Qy   1742 aaatgagtaaataaatgttccatgtaaagtacacgctgaagatgacataagacagaaattta 1801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35610 aaatgagtaatataaatgttccatgttaagtacacgctgaagatgacataagacagaaattta 35669

Qy   1802 cacagctaccctcacataaaacacagaatgcatttcttgacaatttccgaacacacat 1861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35670 cacagctacctctacataaaacacagaatgcatttcttgacaatttccgaacacacat 35729

Qy   1862 taagtgaaaatgacttagaagaactaaggtagtcactaataatgtaacacatacaggcat 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35730 taagtgaaaaatgacttagaagaactaaggtagatcactataaaaatgtaacataacaggcat 35789

Qy   1922 ctgtacatgtttctgattccgtacagataatagtggatctcaacacaacagaagtcag 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35790 ctgtacatgtttctgatttcagtacagataatatggatctcaacaaacagaagtcag 35849

Qy   1982 atactgtgctttccacgcaaaggaaaatgaccttcattcaatattactc 2041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35850 atactgtgctttttccacggcaaaggaaaatggaccttcattcaattattactc 35909

Qy   2042 atgattctggtctgataacaatcaatcccacaagagcaacctaactgttcaggcaaagg 2101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35910 atgattctggtctctgataacaatcacagttccacacacagttcacaggcaaaag 35969

Qy   2102 ctccattccataccctcctcctgaggacccaatgaattgtgacttcaagaaatatggatagtt 2161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  35970 ctccattccaaccttctttactacagctcttacttcaagaatatggatagtt 36029

Qy   2162 taccttctgtaaaatacatcgaaaagtgaaaataattaagacgaaatagaaaagaa 2221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36030 tacctttcttggtaaaatacatcgaaaagtgaaaataataatagaaaaagaaa 36089

Qy   2222 atcctggaaccaaatgctgaaattgataaaaagaaactgaaatttattacaacagaagaaaaca 2281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36090 atctggaaccaaatgctgaatttgataaaaagaactgaaatttattacaacagaagaaaaaca 36149

Qy   2282 gaattgtagttcaccggtacagtcttcactagagtgaagaggaat 2341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36150 gaattgtagttcaccggtacagttctttactgacttgtggaagagaat 36209

Qy   2342 cagaaatttgggttcacaacgctacacagaaagaagtggtatatgcaaatgttttagata 2401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36210 cagaaattttggggttcacagctacacagaaagaagtggtataagcaaatgttttagata 36269

Qy   2402 tttgggaagaagaaaattcagataatctcgttaacacagacgtttttctcgtcccctcaactt 2461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36270 tttgggaagaggaagaaaattcagataatctcgtTAACACCGTTTTCTCGTCCCTTCAACTT 36329
```

Right column

```
Qy   2462 ctacattactggcttttaggaatttaaaaaaatgcaatactttctcaagaagtgataaggatca 2521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36330 CTACATTACTGGCTTTAGGAATTTATAATATGTTTAGAATTTAAAAAAATGCAATACTTTCTCAGAAGTGATAAGGATCA 36389

Qy   2522 tattcctgaaattttttataaatatgtatggaaattcttaggatttttttaccagctttgt 2581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36390 TATTCCTGAAATTTTTTATAAATATGTATGGAAATTCTTAGGATTTTTTTTTACCAGCTTTGT 36449

Qy   2582 ttacagaccaaatgtaaaatattaaaatattgcaattttcacagaaattgaata 2641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  36450 TTACAGACCCAAATGTAAAATATTAAAATATTTGCAATTTTCACAGAAATTGAATA 36509

Qy   2642 cctgttaaaagaaaaattacacagaataaactgtgactggtctgttttttacacattaaaaaa 2700
          ||| |||||||||||||||||||||||||||||||||||||||||| ||| ||
Db  36510 CCTGTTAAAAGAAAAATTACACAGAATAAACTTGTGACTGGTCTTGTTTTTTACATTATTATATA 36568


                               RESULT 9

LOCUS       AC004958   163889 bp   DNA             HTG       12-JUN-1998
DEFINITION  Homo sapiens clone DJ1096F22, *** SEQUENCING IN PROGRESS ***, 6
            unordered pieces.
ACCESSION   AC004958
VERSION     AC004958.1  GI:3213039
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163889)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163889)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     NOTE: This is a 'working draft' sequence. It currently
            consists of 6 contigs. The true order of the pieces
            is not known and their order in this sequence record is
            arbitrary. Gaps between the contigs are represented as
            runs of N, but the exact sizes of the gaps are unknown.
            This record will be updated when the finished sequence
            as soon as it is available and the accession number will
            be preserved.

FEATURES             Location/Qualifiers
     source          1..163889
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DJ1096F22"
                          1..163889
                     1         1462: contig of 1462 bp in length
                     1463      1480: gap of unknown length
                     1481      3064: contig of 1584 bp in length
                     3065      3082: gap of unknown length
                     3083      4710: contig of 1628 bp in length
                     4711      4728: gap of unknown length
                     4729      6530: contig of 1802 bp in length
                     6531      6547: gap of unknown length
                     6548     16429: contig of 9882 bp in length
                    16430     16446: gap of unknown length
                    16447    163889: contig of 147443 bp in length.

BASE COUNT     49499 a   30556 c   31570 g   52166 t     88 others
ORIGIN

Query Match      43.9%;   Score 1194.2;   DB 2;   Length 163889;
Best Local Similarity   99.7%;   Pred. No. 2.1e-216;
Matches 1196;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

                    /clone="DJ1096F22"

Qy   1502 aaagaataaatacagtgttggatccttctcctgtttctgcaagtgtcctgaaaaga 1561
```

Db 74946 AAAGAATAAAATACAGTGTTGGATCCCTTTCTCCTGTTTCTGCAAGTGTCTCCTGAAAAGA 75005

Qy 1562 ctgaacaaaagagaaaagtggaattgcaacatcatattctcagagaagattgccaggaagtg 1621
Db 75006 CTGAACAAAAGAGAAAAGTGGAATTGCAACATCATATTCTCAGAAAAGATTGCCAGGATG 75065

Qy 1622 ataccaacagtgaaggagcagaattccctgtataaagagagacccaggaaactgaaaaaagc 1681
Db 75066 ATACAACAGTGAAGGAGCAGAGAATTTCCGTATAAAGAGAGACCCAGGAAACTGAAAAAGC 75125

Qy 1682 tcctgttttattcagagaccccatcccccaccccttcaaatgaatctgaagtgggcttaatgaga 1741
Db 75126 TCCTGTTTTATTCAGAGACCCCATCCCCCACCCCTTCAAATGAATCTGAAGTGGGCTTAATGACA 75185

Qy 1742 aaatgagtaataaatgttccatgttaagtacagctgaagatgactcataagacagaatttta 1801
Db 75186 AAATGAGTAATAAAATGTTCCATGTTAAGTACACAGCTGAAGATGACTACAGAATTTTA 75245

Qy 1802 cacagctaacctctcacataaaaaacacaggaatgcaattcctgacattccgaacacat 1861
Db 75246 CACAGCTACCTCCTACATAAAAAACACAGGAATGCATTCGTTGACATTTCCGAACACACAT 75305

Qy 1862 taagtgaaaatgacttagaagaactaaggtagatcactataaatgttaacatacaggcat 1921
Db 75306 TAAGTGAAAATGACTTAGAAGAACTAAGGTAGATCACTATAAATGTAACATACAGGCAT 75365

Qy 1922 ctgtacacatgtttctgatttcagttacagataatagtgttgatctcaaccaaacacaaagtcag 1981
Db 75366 CTGTACACATGTTTCTGATTTCAGTTACAGATAATAGTGGGATCTCAACCAAAACACAGTCAG 75425

Qy 1982 atactgttgcttttctccacaaaaggaatctcaaggaaaagggaaagg 2041
Db 75426 ATACTGTGCTTTTCCACGAAAAGGATCTCAAGGAAAAGGGACCTCAAATATTTACTC 75485

Qy 2042 atgattctggttctgaataacaataaacagttcacaagagacacctcaactgttcaggaaag 2101
Db 75486 ATGATTCTGGTCTGATAACAATAAAACAGTTCACAAGAGACCCTAACTGTTCAGGCAAAGG 75545

Qy 2102 ctccaattccaatactcctcctgaggaaccaatgaattgtgacttcaagaataatggtagtt 2161
Db 75546 CTCCATTCCATAACTCCTCTGAGGAACCAATGAATTGTGACTTCAAGAATATGGATAGTT 75605

Qy 2162 tacctcttggtaaaatacatcgaaaagtgaaaataataataggacgaaaatagaaaaagaaa 2221
Db 75606 TACCTTCTGGTAAAATACATCGAAAAGTGAAAATAATAATTAGGACGAAAATAGAAAAGAAA 75665

Qy 2222 atctggaaccaaatgctgaattgataaagaagaactgaattctatcacacaagaagagaaaaca 2281
Db 75666 ATCTGGAACCAAATGCTGAAATTTGATAAAGAAGAACTGAATTTATTACACAAGAAGAAAACA 75725

Qy 2282 gaattgtagttcaccggtacagtctcttactagactgttctcagagtcagtgaagagaaat 2341
Db 75726 GAATTGTAGTTCACCGGTACAGTCTTTACTAGACTTGTTCTCAGAGTCAGTGAAGAGAAAT 75785

Qy 2342 cagaaattcttgggttcacaagctaacacaagctacacagagaaaaagaaaagttatatgcaatgtttagata 2401
Db 75786 CAGAAATTTGGGTTCACAAGCTAACACAAGAAAAAGAGTGGTATATGCAATGTTTTAGATA 75845

Qy 2402 ttgggaggaaggaaattcagaataaatctgttaacagcgtgtttctcgtcccctttcaactt 2461
Db 75846 TTTGGGAAGGAAGGAGAAAATTCGTTAACAGCGTTTTTCTCGTCCCTTCAACTT 75905

Qy 2462 ctacattaactggcttttagaattcgcataatgtcatacactttcctgagaagtgatataggatca 2521
Db 75906 CTACATTTACTGGCTTTTAGAATTCATAACTTTTCACAAAGTGATAAGGATCA 75965

Qy 2522 tattcttgaaattttttatatataaaaaatggaaattttaccagcttttgt 2581
Db 75966 TATTCTTGAAATTTTTATATAAAATGTATGGAAAATTCATAACTTTTACCAGCTTTGT 76025

Qy 2582 ttacagaccccaaatgtaaaatattaaaaatatattgcaattctctacagaattgaata 2641
Db 76026 TTACAGACCCCAAATGTAAAATATTAAAATAAATTGCAATTTTCTCTACAGAATTGAATA 76085

Qy 2642 cctgtaaagagaaaattacagaataaacttgtgactgtcttgttttacactaaaaaa 2700
Db 76086 CCTGTTAAAGAGAAAATTACAGAATAAACTTGTGACTGGTCTTGTTGTTTTACATTATATATA 76144

Qy 445 ggaagccgtgctttcgggctgcgccggtgcgaacacttctccggaccacgtaggtg 504
Db 22 GGGAGCTTTGTGTGTTCGGCGGCCGTTCCAAGGGCGCCCCGGTGGGACCCAGCAATGTAGGTG 81

Qy 505 ccgggcgactgccatagaacctccggaggaagaatctaaaggtacaattccaaggg 564
Db 82 CCTGGTGGCTGCCATGAACCTCCGGAGGACCATGAGGATCCACCAGCAAAAGCCACTCTCTCCCGGG 141

Qy 565 tggaatccaagtcaaaaatgaaaaaaaacaacagaaccatcctgaaatctgaaaactgataa 624
Db ... tggaatccaagtcaaaaatgaaaaaaaacaacagaaccatcctgaaatctgaaaactgataa

```
Db   142  TGGCATCCAAGACAGAAAATGAAAAATCGACCATCCTTGAAATCCCTGAAGGCCGATAA  201
          |||  ||||  ||||| ||||||||||| ||||||||||||| || |||||||||||||
Qy   625  caggcagaaaatccaaatgtaagcacttgggaaaagtattacttgacttacc        684

Db   202  TAGACTGGAAAAAATCCAAATATAAGCCCTTTGGGAAAATATTTACCTTGACTTGCC    261
          |||  ||| |||| |||||| || |||| || ||||||||  || ||||| |||
Qy   685  ttctgtcacccatatctgaaaaacttcaaaaggacaattaaggatctgggaggccagttga  744

Db   262  TTCTATCACCATATGTGAGAAAAACTACAAAAAGATATAAAGGAAACTCGGAGGCCGAGTTGA  321
          ||| ||||||||||||||||| || ||||||| |||||||||||||||||||||||||||||
Qy   745  agaattcctcagcaaagatcagtgttatctattcatttcaaataagaggaagctaaattgc  804

Db   322  AGAAATTTCTCAGCAAAGATAATCAGTTACTTTGTTTCAAATAAAAGGAAGCAAAGTATGC   381
          |||  ||||||||||||||| ||||||| ||| ||||||||||  |||||||| |||| ||
Qy   805  acaaaaccttgggtcgaattcctcctgtaccaagtccagatatctgatatactgcagaaac  864

Db   382  ACACAACATTGGGTCGCGGGGTTTCTCTGTACCAGTCCAGAAATCTGCATATACTCCAGAAAAC  441
          ||  |||||||||| | |  |||| ||||||||||||||||| |||| ||||||| ||||| |
Qy   865  cacttcacctcatccaagccaatgtggaagttcatttaagtcaacagagacacagtgtgtt   924

Db   442  CACCTTCTCCATCCCACAGCGAAGTTCCTTAAGTCTCAAGACACAGTATGTTT          501
          ||| || | ||| || ||||||||| || ||||| ||||| ||||| || |||
Qy   925  aagcagaggaaaattattagttgaaaaagctatccaaggacaatgaatcaagttattccttcaaa  984

Db   502  GAGCCAGGGCGAAAAATTGTTAGCTAGAAAAAGCTCGTCGAAGGACATCATGAATTTTATTCCTGCTAA  561
          |||  || | ||||||  || |||| |||||| | |   ||||| ||| ||||| ||| || ||||
Qy   985  tagtatattattcaaatgcctttgtcatgggagtaaaaaaatttcttcatattgatgacattg  1044

Db   562  TAGTATTATTATTCAAATGCCTTTATCATTGGGGAGTAAAAGAATTCTTCATATTGACATTCG  621
          ||||||||||||||||||||||||| ||| |||||||||| |||| |||||||||||| ||| |
Qy  1045  atactacattgaacaaagaagaaaacaagaactcaagaaaatcaagtactactccagt  1104

Db   622  ATACTACATTGAACAGAAGAAGAAAACACTCGAGTGCACTCAAGAAAATCCAGTGCTTCTCG  681
          ||||||||||||||| |||| ||||||||| |||||| ||||||||||||||||||  ||  |
Qy  1105  aagagatggggggcaaaagagttggttagttggtgcacaaaaaacaagaacagaagctca  1164

Db   682  GAAGGAGATGCGGGGGCAAAAAAGCTCGGCCCCTGGAATACAAGAACAGGGAGACTCAA    741
          ||||||| || ||||||||||   | ||   |||| |||||||||| ||  |||| |||
Qy  1165  aaaagccttttgtaaagtgggagagtatgtagcc                          1196

Db   742  GAAAGCCTTTTTTAAAAGGTTGAAGAGTGAAACGGTGCTACAGGCCCATTTTACCTTCAGCT  801
          ||||||||| | |||| ||||||| || || | ||||  ||||||||| ||| |||||||
Qy  1197  ---------------aagccctgcagtccacttgatgtaga                    1223

Db   802  CCCCAGCTTCTCCGGTGTATAAVTACTTTCTTCACGAAGCCCTGCGAGTCCGTTTGATATAGA  861
          |||||||| || ||||  | |  ||||  ||||| |||||||||||| || |||||||||||
Qy  1224  caagccaatctagttatgcaaaggcaaactcaggttaaactaagaatccaaacagatggcga  1283

Db   862 ·GAAGTCATCCAGGGGTCAGAAACACAACCTCAGAAAACCTCAGGAATTTAATATGCGACGGCGA  921
          ||||||||||||||| |||||| | |||| || ||||  |||||||  ||||| ||| ||| |
Qy  1284  taagtatgtgtgaacctcaatcactccagttgaaagaagagaaaagagatatg          1343

Db   922  TAAAT--------GTGGAACCCAGTTCAAACTCCACAGTTCTGAACTCACTCTGAGTGAGAAACACAG  978
          |||||        |||| |  ||  | |||||| | ||  |||| ||| ||  |||| ||||||
Qy  1344  tgaatgttgctgcagaagtactcaagttgtgtgatgatatctaagttagttt          1403

Db   979  TGAATGTTGCTTACGAGTTCAAACTCAGGAGGAAAAACTCGAACGAAAAAGGATATTG  1038
          |||||||||||   | |||||| ||||||||| ||| || || | ||||||| |||||
Qy  1404  aaaacttgcacgagtaaccagtatcaaagttgtgatgtatctaagttagttt        1463

Db  1039  GAACTTTGCACAGAGTAACCAGTATCAAGTTGTTGCATGTAATGTGTAAGTTAGTTGT  1098
          |||||||||||| |||||| |||| |||| |||| |||| ||||||| |||||||||
Qy  1464  tgacttgttggaatatggaaaggacacacctaaaagaaagaataaaatacagtgttgg  1523

Db  1099  CGATTTTGTGTGGAAATAATGGAGAGACACACAGTTCAAAAGAATAACAGTGTTCGG  1158
          ||||||||| |||| |||| |||| || |||||    |||  |||||  ||||||
Qy  1524  atcccttcctcctgtttctgaaaagctgaaccaaaaagaaaagtgga             1583

Db  1159  GTCCCTTTCTTCCTGTTTCTCGCAAATGTCCTGAAAAACACTGACCACGAAGGACAGCCACT  1218
          |||||||||||||||||| |  |||| |||||| ||||||| ||| | ||  ||  |||||
Qy  1584  attgcaacatattctcagaaagaatgcccaggagaatgatacaacagctgaaggagacagaa  1643

Db  1219  CTTGGACCCCAATTTCCAGAAGGACGTGGGCGCAAAGCACCTGCTCAAGCCGAA  1278
```

```
Qy  1644  tttcctgtataagagacccaggaaactgaaaaaaagtctcctgtttattcagagccat  1703
          ||| ||||||||||| |||| | ||||||||||||||||| |||| ||||| ||| |||
Db  1279  CTCTCAGTATGAAGACCAGAGAAAACCCAGAAAACCTGAAGAAAAAGCATGGGGTTTCAGAGCCCAC  1338

Qy  1704  ccccaccccttcaaatgaattgagaggggcttatgagaaatgagtaataaaatgttccat  1763
          |||||||||||||||||| | |||||||  |||||||||||| ||| |||||||||||||
Db  1339  AACCTATTCTTCAGCTGGATTGAAAGGGATGTGA-------CAGGAAACCAGTGAGCAT  1389

Qy  1764  gttaagtacagctcgaaagatgacaataagacagacagaaatttcacagactacctctacataaaa  1823
          || ||||||||||| |||||||||| |||| || || |||||| |||  |||| |||||||||
Db  1390  GTTCAATGCAAGTGCAAGCCTGACCCACAGACAGAGTCAGTTAGTCACACAGCTCACCTCTCCACATTAAAA  1449

Qy  1824  caaacaggaatgcattcttcgaacacacattaagtgaaaatgacttagaaga  1883
          || |||| ||||| | |||| ||   |||  ||| ||||| |||||  ||||
Db  1450  CACACCGGAAC--------------ACCAAGTGACCAGGGTGTAGAAATGAC---GGAGA  1491

Qy  1884  actaagggtagatcaactataaatgtaacatacaggcatcgtgtacacagtgtttcctgattcag  1943
          |||||| |||| ||||| ||||||||| || | ||||| || |||| ||||| ||| ||||||||
Db  1492  ACAGACAGTAGACCCCTGCTCCAGGCGTGTCCCAGTCATGTGGGACAACAAGTTTCTCATCTCAG  1551

Qy  1944  tacagataatagttggatctcaaccaaacagaagtcaagataactgtgctttttccagcaaa  2003
          |||||||||||| || | | |||||||| || || |||||||||| | ||||||||||||||
Db  1552  CACAGAAAGTAATCTCACCCCAACCACAGTTAGCTCGACAAGTTAGCTGTCACAATTACACAGCTTTCAGCAAA  1611

Qy  2004  ggattcaaggaaaagggaccttcaattcaatccaatattcttacttcttcatgtctgataacaat  2063
          |||||||| | ||||||| || || | ||||| |||||| |||  |||| |||||||||| ||||
Db  1612  AGAATCTGCAAGAAAAGGAGCCTTCAATGTCATGTTGTGTATAGGCGGACTGGTCATGTTGACCCTGGCATT  1671

Qy  2064  aaacagttcacaagagcaactaactgttcaggcaaaggctccattccaatctcctctga  2123
          |||||||||||||||| ||| |||||| || || |||||| ||||  | ||| | |||
Db  1672  AAATACATCAAAAGAGCACGTAACAATGAAAGGCGACGACAACTCACCACCCTCTAGTCCC---A  1728

Qy  2124  ggaaccaaatgaaatgtgacttcaagaatatggataagttcacctctcgttaaaatacaatcg  2183
          ||||||||||  ||  ||||| ||||||||||| ||||| ||||| || |||| |||| |||
Db  1729  GGAGCCTGTCAATGCAGTGTCGACACTGGACAACAATTTGCCCCTGTGTAAAATCCAGCG  1788

Qy  2184  aaaagtgaaaataatattaggacgaaagtagaaacaaatctgaacacaaatgctgaatt  2243
          ||||  || |||| ||| || | || |||| |||||  |||  | | |||||||| ||
Db  1789  GAAAGTGAGAATGCTATTAGGGACAACA---GAAAGCAAAATGCGAACCGAGTTGCTGAAATT  1845

Qy  2244  tgataaa------agaactgaattattaccacaagaaagaaaaacaaattgttagttcaaccggt  2300
          || |||       | | | |||| ||||||| |||| |||| | || | | | |||| ||||||
Db  1846  GGATAAAAGATGAAAAAGAGAACTGAGTATTCTCCCTGCACATGAAGGAAGACAAGAACTTCGTGGGTCACCAGT  1905

Qy  2301  acagtctcttactagacttgtcttcagactagtgaagagaagaactcagaatctttgggttcac  2360
          |||| | ||||| | ||| | |||| | | |  ||||| |||||| |||| |||| || | |||
Db  1906  ACAGTCTCTATTGGACCTTTTTCAGAAGAGCGAGAGAAATCAGAAATTTTTGGGTTTTAC  1965

Qy  2361  aagctacacagaaaagagtggtatatgcaatgtttagtatattgggaaggaggaaaatc  2420
          |||||||||||||||| | ||| ||  |||||| |||||  ||||||||| |||| ||||
Db  1966  AGGCTACACTGAAAACAGTGGAATAATGTGCAATGGTTTGCATGGTTTTTAGATANTTTGGAGGAAGAAAAAATTC  2025

Qy  2421  agataaatctgttaacagcgcttttctctgtcccccttcaacttcaacttactggctttta  2480
          || ||||| || |||||| | ||||  | ||| ||||||||| |||||| ||| |  ||||
Db  2026  AAGTACCCTGTTGCTAACAACTTCTGCAACTTCTCGTTTTTGTTGGAGAGAAGAAGAAAAATTC  2085

Qy  2481  gaatttaaaaaatgcatactttcagaagttgaaggtgatggaatcatattcttgaaatttttata  2540
          |||||| || ||  |  |||||||||| || ||||| || |||||||||| | || |||| ||
Db  2086  GTATTAAAANGTGCTTNTCAGAAGTTGACAGAGCTATATTCTCCAGAGTTTTNTAATAACATAT  2145

Qy  2541  aatatgtatggaaattcttaggattttttggttttcttttcagcagccaaagtgaaa  2600
          |  ||| |||||||| ||||| ||| |||||| |||||||| |||| ||| |||
Db  2146  ATAATTACTGGTGGGGGGTTTTGGTTTTTNTGTTTTTGGCCCACACTTTTNTNTTAGGGACTGTGTAA  2205

Qy  2601  tattaa--aaataaatatttgcaattttctacagaattgaatacctgttaagagaaaaatt  2658
          | ||||  | | | | ||||| | | ||||| |||||  | | || ||||| | |||| ||
Db  2206  TATTAATTAAGTGGTGTTTGTGTTTTNTTTTNTNTTACAGGATTTGACTTAGCTGTT----ACAAATT  2261

Qy  2659  acagaataaacttgtgactcggtctgttgttttcaattaaaaaaaaaaaa  2706
          ||||||||||| || |||  | ||  ||| |  || |||||||||||||
Db  2262  ATACAATAAAAAACTTAATGACTTAATTTATTTTTATAAAAAAAAAAAAAAA  2309
```

```
RESULT 11
LOCUS       AK022969        2276 bp      mRNA            PRI     29-SEP-2000
DEFINITION  Homo sapiens cDNA FLJ12907 fis, clone NT2RP2004396, highly similar
            to Homo sapiens mRNA for activator of S phase kinase.
ACCESSION   AK022969
VERSION     AK022969.1  GI:10434668
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,
            clone_lib:NT2RP2 clone:NT2RP2004396.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,
            Matsunawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
            Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Sasaki,N.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished (2000)
REFERENCE   2  (bases 1 to 2276)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
            Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:genomes@hri.co.jp,
            Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing; Research Association for Biotechnology; cDNA library
            construction; 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1..2276
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NT2"
                     /cell_type="teratocarcinoma"
                     /clone="NT2RP2004396"
                     /clone_lib="NT2RP2"
                     /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
                     precursor cells after 2-weeks retinoic acid (RA)
                     induction."
BASE COUNT      793 a    381 c    421 g    681 t
ORIGIN

Query Match             40.5%;  Score 1101.6;  DB 9;  Length 2276;
  Best Local Similarity   99.6%;  Pred. No.6.2e-199;
  Matches 1104;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1593 tatttctcagaaagatgccaggaagatgatacaacagtgaaggaggagcagaattcctgta 1652
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         1 TATTTCTCAGAAAGATTGCCAGGAAGATGATACAACAGTGAAGGAGGAGCAGAATTCCTGTA 60

Qy      1653 taaagagaccaggaaactgaaaaaaaagtcctcgtgtttattcaggagcccatccccaccc 1712
             |||||||||| |||||||||||||||||| ||||||||||||||| ||||||||||||||||
Db        61 TAAAGAGACCCAGGAAACTGAAAAAAAAGCTCTCGTGTTTATTCAGAGCCCATCCCCACCC 120

Qy      1713 tccaatgaattgagagggcttaatgaataataaatgttccaagtcatgttaagtac 1772
             |||||||||||| ||||||||||||||||| ||||||||||| |||||||||||||
Db       121 TTCAAATGAATTCAGAGGGCTTAATGAATAAATGTTCCATGTTAAGTAC 180

Qy      1773 agctgaagtgacatagacacagaaatttcacacagctacctctacataaaacaaacagga 1832
             ||||||| |||||| ||||||||||||| ||||||||||||||||||||| ||||||||||
Db       181 AGCTGAAGATGACATAAGACACAGAAATTTTACACAGCTACCTCTACATAAAAACAAACAGGA 240

Qy      1833 atgcaattcttgacaatttccgaacacacaattaagtgaaaatgacttagaagaactaaggt 1892
             ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||| |
Db       241 ATGCAATTCTTGACAATTTCCGAACACACATTAAGTGAAAATGACTTAGAAGAACTAAGGGT 300

Qy      1893 agatcactataaatgtaacatacacaggcatcctgtacatgttcctgtatttcagtacagataa 1952
             ||||||||||||||||||||||||||| ||||| |||||||||| |||||||||||||||||||
Db       301 AGATCACTATAAAATGTAACATACACGGCCATCTGTACATGTTCTCGTGATTTCAGTACAGATAA 360

Qy      1953 tagtggatctcaaccaaacaaacagaagtcagaatactggctgacttttccagcaaaggatctcaa 2012
             |||| ||||||||| |||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TAGTGGATCTCAAACCAAACCAAACAGAAGTCAGAATACTGGCTGACTTTTCCAGCAAAGGATCTCAA 420

Qy      2013 ggaaaaggacgcttcattcaatattcatctactcatgatttgtctgataacacaaacagttc 2072
             |||||||||| ||||| |||||||||||||||||||||| |||||||||||| |||||||||||
Db       421 GGAAAAGGACCCTTCAATTCAATATTCATCTACTCATGATTCTGGTCTGATAACAATAAACAGTTC 480

Qy      2073 acaagaggcacctaactgttcaggcaaaggcttccatccaatccactcctctggagaacccaa 2132
             |||||| ||||||||||||||||||||||||||||||| ||||||||| ||||||| |||||||
Db       481 ACAAGAGCACCTAACTGTTCAGGCAAAGGCTTCCATTCCAATCCACTACTCCCTCTGGAGGAACCCAA 540

Qy      2133 tgaatgtgacttcaagaatatggatagtttacctctggtaaaatacacatcagaaagtgaa 2192
             |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db       541 TGAATGTGACTTCAAGAATATGGATAGTTTACCCTCTGGTAAAATACACATCAGAAAGTGAA 600

Qy      2193 aataatattaggacgaaataagaaaaatctggaaccaaagtgctgaattgataaag 2252
             ||||||||||||||| ||| |||||||||||||| ||||||||||||| ||||||||
Db       601 AATAATATTAGGACGAAAATAGAAAAATCTGGAAACCAAAGTGCTGAATTTGATAAAG 660

Qy      2253 aactgaatttattacacagaagaagaaacacagaatttgtagttcaccggtacagtcttact 2312
             ||||||||||||||||| ||||||| ||||||||||||||||||||||  ||||||||| |||
Db       661 AACTGAATTTATTACACAAGAAGAAAAACACAGAATTTGTAGTTCACAAGGTACAGTCTTTACT 720

Qy      2313 agagctgtttcagactagtgaagagaaatcagaaatttttgggtttcacaagctacacaga 2372
             |||||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       721 AGAGCTTGTTTCAGACTAGTGAAGAGAAAATCAGAAATTTTTGGGTTTCACAAGCTACACAGA 780

Qy      2373 aaagagtggtatatgcaatgtgtttagatattgtttggaagaagaaattcagataatctgtt 2432
             |||||||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||||
Db       781 AAAGAGTGGTATATGCAAATGTGTTTAGATATTGTTTGGAAGGAGAAATTCAGATAATCTGTT 840

Qy      2433 aacagcgtttctctcgtccccttcaacttctacattactggctcttactgtcagtcttttt 2492
             |||||||||| |||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db       841 AACAGCGTTTTCTCTCGTCCCCTTCAACTTCTACATTACTGGCTTTTACTGTCAGTCTTTTT 900

Qy      2493 tgcatactttctcagaagtgataaaggatcatctttctggaaatttctttatataaatatggga 2552
             |||||||||| |||||||||| ||||||||| |||||| |||||||||||||| |||| ||||||
Db       901 TGCATACTTTTCAGAAGTGATAAGGATCATATTTCTTGAAATTTCTTTATAAATATGTATGGA 960

Qy      2553 aatttcttaggtgtttttttttacccagctcttgttctacggaccccaaatgtaaatattacaataa 2612
             |||| ||||||||||||||| |||| |||||||||||||||||||||||||||||||||| |||||||
Db       961 AATTCTTAGGTGTTTTTTTTACCAGCTCTTGTTCTACGGACCCCAAATGTAAATATTAAAAATAA 1020

Qy      2613 atattgcaatttctacagaattgaatacctgttaaagagaaaaatttacaagaataaaacttg 2672
             |||||||||||||||||||||||||||||||||||||||||||| |||||| |||||||| |||
Db      1021 ATATTGCAATTTCTACAGAATTGAATACCTGTTAAAGAGAAAAATTTACAGAATAAAAACTTG 1080

Qy      2673 tgactggttctctgtttttacattataaaaaa 2700
             |||||||||||| |||||||||||||| | | |
Db      1081 TGACTGGTTCTCTTGTTTTTACATTATATATA 1108


RESULT 12
LOCUS       AC091341        232903 bp      DNA             HTG     18-APR-2001
DEFINITION  Rattus norvegicus clone CH230-1C19, WORKING DRAFT SEQUENCE, 35
            unordered pieces.
ACCESSION   AC091341
VERSION     AC091341.1  GI:13661907
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 232903)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
```

```
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,J., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Louseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.
            and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 232903)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-APR-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
              Center project name: TUCU
              Center clone name: CH230-1C19
            -------------- Summary Statistics
              Sequencing vector: Plasmid; M77789
              Sequencing vector: M13; L08821
              Chemistry: Dye-primer Bodipy: 33% of reads
              Chemistry: Dye-terminator Big Dye: 67% of reads
              Assembly program: Phrap; version 0.990329
              Consensus quality: 230627 bases at least Q40
              Consensus quality: 245439 bases at least Q30
              Consensus quality: 245611 bases at least Q20
              Consensus quality: 253380 bases at least Q20
              Estimated insert size: 222874; sum-of-contigs estimation
              Quality coverage: 0x in Q20 bases; agarose-fp estimation
              Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 35 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1     28075: contig of 28075 bp in length
            *      28076     28175: gap of unknown length
            *      28176     42563: contig of 14388 bp in length
            *      42564     42663: gap of unknown length
            *      42664     62432: contig of 19769 bp in length
            *      62433     62532: gap of unknown length
            *      62533     79400: contig of 16868 bp in length
            *      79401     79500: gap of unknown length
            *      79501     92855: contig of 13355 bp in length
            *      92856     92955: gap of unknown length
            *      92956    103757: contig of 10802 bp in length
            *     103758    103857: gap of unknown length
            *     103858    117059: contig of 13202 bp in length
            *     117060    117159: gap of unknown length
            *     117160    126002: contig of 8843 bp in length
            *     126003    126102: gap of unknown length
            *     126103    134408: contig of 8306 bp in length
            *     134409    134508: gap of unknown length
            *     134509    144061: contig of 9553 bp in length
            *     144062    144161: gap of unknown length
            *     144162    153305: contig of 9144 bp in length
            *     153306    153405: gap of unknown length
            *     153406    161399: contig of 7994 bp in length
            *     161400    161499: gap of unknown length
            *     161500    168278: contig of 6779 bp in length
            *     168279    168378: gap of unknown length
            *     168379    175475: contig of 7097 bp in length
            *     175476    175575: gap of unknown length
            *     175576    180952: contig of 5377 bp in length
            *     180953    181052: gap of unknown length
            *     181053    185776: contig of 4724 bp in length
            *     185777    185876: gap of unknown length
            *     185877    189454: contig of 3578 bp in length
            *     189455    189554: gap of unknown length
            *     189555    193485: contig of 3931 bp in length
            *     193486    193585: gap of unknown length
            *     193586    198298: contig of 4713 bp in length
            *     198299    198398: gap of unknown length
            *     198399    201470: contig of 3072 bp in length
            *     201471    201570: gap of unknown length
            *     201571    205298: contig of 3728 bp in length
            *     205299    205398: gap of unknown length
            *     205399    208583: contig of 3185 bp in length
            *     208584    208683: gap of unknown length
            *     208684    211371: contig of 2688 bp in length
            *     211372    211471: gap of unknown length
            *     211472    215483: contig of 4012 bp in length
            *     215484    215583: gap of unknown length
            *     215584    217293: contig of 1710 bp in length
            *     217294    217393: gap of unknown length
            *     217394    218893: contig of 1501 bp in length
            *     218894    218994: gap of unknown length
            *     218995    220090: contig of 1796 bp in length
            *     220091    220790: gap of unknown length
            *     220791    222890: contig of 2263 bp in length
            *     222890    223153: gap of unknown length
            *     223154    224510: contig of 1257 bp in length
            *     224511    224610: gap of unknown length
            *     224611    225959: contig of 1349 bp in length
            *     225960    226059: gap of unknown length
            *     226060    227888: contig of 1829 bp in length
            *     227889    227988: gap of unknown length
            *     227988    229034: contig of 1046 bp in length
            *     229035    229134: gap of unknown length
            *     229135    230482: contig of 1348 bp in length
            *     230483    230582: gap of unknown length
            *     230583    231761: contig of 1179 bp in length
            *     231762    231861: gap of unknown length
            *     231862    232903: contig of 1042 bp in length.
```

```
FEATURES          Location/Qualifiers
source            1..232903
                  /organism="Rattus norvegicus"
                  /db_xref="taxon:10116"
                  /clone="CH230-1C19"
BASE COUNT    66408 a   49441 c   48322 g   65273 t   3459 others
ORIGIN

Query Match     33.1%;   Score 900.6;   DB 2;   Length 232903;
Best Local Similarity  68.7%;   Pred. No. 9.5e-161;
Matches 1627;   Conservative 0;   Mismatches 559;   Indels 181;   Gaps 21;
```

```
Qy    420 gcggcgccgtcctgtcaacaggccggggggaagccgtgcttcgcggctgcgccggttgcgaacac  479
          ||||||  |||||||||  |||||||   |||||||||||||||| | |||||||||||||||||
Db  39488 GCCGAGCCAGGACTGGTGGGCGGGCCCTTGGAGCTTTGTGTTTCGCGGGCTGTCCCAGTGGCTC 39429

Qy    480 tttctccggacccagcatgtaggtgccggggactgccagtgaactccggaacgccatgagga  539
          |||| |||| ||| |||||||||| | | |    |||| | | |||  |||  ||||||||||
Db  39428 CCGGTGCGGGACTCAGCATGTAGGTAGTGGCCTGGTGGCTGCCTGCCTCAGCACCATGAGGA 39369

Qy    540 tccacagtaaaggacatttccaggggtggaatccaagtcaaaaaatgaaaacagacat  599
          ||||||| ||| |||| |||||  ||||| |||||| || |  ||||||||||| |||
Db  39368 TCCACAGCCAAGCACACTTCCTGGGTGGCATCCAAGAAAATGAAGAAATAGACCAT 39309

Qy    600 ctctgaaatctctgaaaactgataacaggccagaaaaaaatccaaccactttggg  659
          ||| |||||||||||| |||||| |  ||||| ||||  ||| | |  |||||||
Db  39308 CTTTGAAAATCCCTGAAAAACTGATAACAGACTGGCATCCAAGAACAAAATATAAAGCACTTTGGG 39249

Qy    660 gaaaagtatttaccttgacttacctctgtcaccatatctgaaaaacttcaaaaggaca  719
          |||| || |||| ||||| |||||    | ||||  |||||||||||  |||||| |||
Db  39248 AGAAAATATCTTACCTTGACTTGCCTTATCTACCAIATGTCAAAAGCTACAAAAAAAGACA 39189

Qy    720 ttaaggatctggggcggcggggttgaaggaatttctcacagacaggccaactcttattt  779
          |||||  ||| |||  |||| ||||||||||||||||   ||    |  ||||| ||||
Db  39188 TAAAGGAACTTGGAGGGCGGAGTTTCTCAGCAAAGTTTATCACTTGTCTTGTTGTTT 39129

Qy    780 caaataaggaaggctaaattgcacaaaaacctttgggtcgaattctcctgtaccaagtc  839
          |||| |||| ||| |||| ||| |||  | |||||| |  |||||| ||| | |||||||
Db  39128 CAAGTAAAAGGAAGAAAAGTAAATGCACAAAACTGGGTCGAGTTCTCCTATACCAAGTC 39069

Qy    840 cagaaatctgcatatactgcagaaaccacttcacctcatccatgatgaagttcat  899
          ||| ||||| ||| |||| || ||  ||| ||||| || ||||||| || ||||
Db  39068 CAGAATCTGCATATACCCGGGAAACCACCTCTCCTCATCCCATGCTTGACAAGTTC- 39009

Qy    900 aggaccatgatttattccttcaaatatagtatattagttgacaaaagctatca  959
          || |||| | || |||  ||||||  || ||| |||| ||  ||||| |||||
Db  39010 ------CACAAGACAGGTTGTGTTTTGAGCCGGAAACCACCTCGCATGCTTGAGAAGTTC---- 38959

Qy    960 aggaccatgatttattcctcaaatagtatatcagtcatgtttgggaagttaa 1019
          | |||   |||||| |||   || |   || |||| |||||||||    |||
Db  38960 -GGATCATGATTTATTCCTGCTAATAGTATAATTATCACGGGAAACTGCCTTATCCCTGTTGGGA 38902

Qy   1020 aaattcttcattgatagatactacattgaacaaaacaaagaaaaaagagttgtatt 1079
          |||||||||||||| | | | ||| || ||  |  || ||| ||| ||| |||||
Db  38901 AGATTCTTCATATTGATGACATTGTGTAAATTGACACAGAAAAGGAAAGCAGGGGTG 38842

Qy   1080 tactcaagaaatcaagtactccagatggggcaaaagagagtcggtagtggtgcac 1139
          |||||||||| | ||| ||||| ||  ||||||| |||||| | | ||||| ||
Db  38841 CACTCAGAAACAAACCAAGCCAGTCTGGGAGGGATCAGGGGTG------AGCTGGTACTGGAATAC 38784

Qy   1140 aaaaaacaagaaggagaccaaaaagccttttgtaaaggtggaggataaatgagagcca - 1197
          |||||  | |||| ||| | ||||  |||||||  |||| | ||||| |||| || ||
Db  38783 AAAAGGCGGAGAACAGGGAGACTCCAGAAGTCTTTTTTAAAAGGTTGAAGATGTGAACCAGA 38724

Qy   1198 ------------------------------------------a 1198
                                                     |
Db  38723 CTTACAAGCTGTTTACCTTCAGCTCCCCAGTATGACGATGACTTTGTATTAACTACTTTCTTCAGA 38664

Qy   1199 agccctgcagtccattgtagaacccatctagtatgtcaaagcaaactcaggtta 1258
          |||| ||| ||| |||| || |  |||||| | |||  |||| ||||| |||
Db  38663 AGCCTTGCCGTCTGTCTG-TATGGATAAGTAATAAATCCAGAAAACAAGGTCAGCCTC 38605
```

```
Qy   1259 aactaagaatccaaacacagatgggcgataagtatggtgggaacctcaattcaactccagttga 1318
          |||| ||| || ||| ||||||| | | || || | |||||  |||||| |||||||||||||
Db  38604 AACTCAGAGAATTAATAATGGACGGCCGGATAAAT---GTGGAACCCCAGTTCAAGTTCCAGTTGA 38548

Qy   1319 aagagaagaagaaaaaaaggatatatgtgaatcgttgcttgcagaaatgaagatctagaaa 1378
          ||| |||||||||||||      |   ||||  |||||||||  ||||||||||| || |
Db  38547 AGGAAAAGAGAAAAAA------------GGAAGTTGCTTGCTCAGAGAATATGAGAGATCTTGA-- 38497

Qy   1379 ctcaccttctaagtgagcaacacagaaactttgcacagagtaaccagtatacaagttgttg 1438
          |||||||||||||| ||||||||||||||  ||||  |||| || ||||||||||||||||
Db  38497 CTCACCTTCTAAGTGCACAACACAGAAAACTTTGCACAGAGTTACCAGTATACAAGTTGTTG 38457

Qy   1439 atgatattgtatctaagttagttttgacttttgtgaaatagaaaggaagcaacactaaa 1498
          |||| | | |||| |||||||||||| ||||| ||| |||||  |||| ||||||||
Db  38457 ACAATATTGTATCTAAGTTAGTTTTGATTTTGTGAAATAGAAAGGAACCAACTAAA 38397

Qy   1499 aagaaaagaatcaaaataacagatgttttggaatcccttttctcctgttttctgaaagtcctgaaaa 1558
          ||||||||||| |||| |||||| ||| | |||| || ||||| || ||||||||| |||||||
Db  38396 GGAAAAGAATAAAGATCACAGTAGTGTTGATGATCCCTTTTTGTGTTTTCTGAAGGTGTCCTGAAAA 38337

Qy   1559 agactgaacaaaggaaaagtggaattgcaacatattttctcagaagatgccaggaag 1618
          || || ||| ||||||  | |||||  |||||||||  |  || ||||  |||| ||
Db  38337 ACTCTGAACAAAGGAAAAGGTACTCTGGTTGGAGCCATGCCCAGTTTCCAGAAGAAAGTCGGAAGAA 38277

Qy   1619 atgatacacaacagtgaaggagaacagaattcttgtaagtacacgtcaggaaaactgaaaaa 1678
          |||| | || | |||||||||||| ||||||||| |||||||||  |  ||| ||| |||||
Db  38277 GCAATGGACAACCAAGAAAAGGAGAAAAGGCTACTGCTTGGAGCCCGTCAAGAAACCTGAAGAAA 38217

Qy   1679 agcctcctgttctttcaggagcccatccccacccttccaaatgaattgagaagggcttaatg 1738
          || ||| ||| ||| ||||| ||| || | ||| |||||| ||||||  || |||| |||
Db  38217 AGCTTGTTGTTTCTTCCAGAGCCCATCAGTTCCAACCCTGCAATGCAAGAACCCGAGTTAAAGGG 38157

Qy   1739 agaaaatgagtaataaatgttccatgttaagtacagctgaagatgacaatgaggaatt 1798
          |||||| ||||| | |||||||| ||||||||||||||| || |  | |||  |||||
Db  38157 AGCTTGTGTTTTGGTTGGAGATGGAGAATGTTCAGAGAATTACCAGTATACCCAACCAGATT 38106

Qy   1799 ttacacagtctaccttcctgataaaaacaaacaagatgcacattcttccgaacaca 1858
          |||||||| ||||| ||| |||| | |||||||| |||||| |||||||| |||||
Db  38106 GTGACAGCATACCTTCTGCATTGGCAATAAACAGGCCATGGCACATGTTCTGAACATA 38046

Qy   1859 cattaagttgaaaatg---acttagaagaaactaaggggtagatcaactatataaatgtaacatac 1915
          ||||||||| ||||||   | ||| ||| | ||  | | || | ||| ||  | ||||| |||
Db  38046 AAATAACTGTAGATGTGGAAGATGATAGAGACCAGGAGGAGTGTCACTCCAACACAAAATTAA 37986

Qy   1916 aggcatctgtacatgttctgatttccagacaataatagtggtacaacagaactaccccaaacaagag 1975
          || ||||||| | |||||||||||| || | ||| || | || ||| ||| |||||| |||
Db  37986 AAATAAACTGTAGATGTGGAAGATGATAGAGACCAGGAGGAGTGTCACTCCAACCAAAATTAA 37926

Qy   1976 agtcagatactgtgctttttcccagacaagatcttcacttcatttt 2035
          | |||| |||| ||  | ||| || |||||||||||| ||| ||
Db  37925 CTGCAGATA------TTACACAGCAAAAGATCTTCTATTTCAGTGTCG 37873

Qy   2036 ttactccatgatctgtctgataacaataacagtccacaagaagcagtccagg 2095
          ||||||| | |||||||||||||| || || |  || ||| ||| || ||||
Db  37872 CTGGTCACATGATTCTGACTGGCAATTAACAGGATTCATCAAGAAGCACCTAACAATGAAGG 37813

Qy   2096 atagttaccttctcggtaaaatacacagtccacagaaagtgaaaatataatattacagtatg 2155
          |||||||| || | ||| | ||| |||| |||| | ||||| | |||| | |||| | |||
Db  37812 CTAGAACTCCACCCTGTAGGTAAAGTGAAAGGTCACAGTGAGATGGTTACACAGTTAGGGCAAAAGG 37756

Qy   2156 caaagtgctccattccatcctcctgagggaggcttcaggaatggtgacttcaagaatatgg 2215
          |||| |||| ||||||| |||| ||||| || | |||| |||| ||| | |||||| |||
Db  37755 AGAGTCTGCCCTTGTCGTAGGATTCCACGGAAAAGTGAAATGATGAGGGTCACAGTGAGATGATATGG 37699

Qy   2216 aagaaaatctggaacaaagctgaattgat---aaaagaactgaaatgctattaccacaag 2272
          || ||||||| ||||| || | ||||||   |||| ||| | |||| |||| ||||||
Db  37698 AAGCCAAGTGTGGAACCTGGTGTCCGGAATAAGTGAGAAGTGAGTATACTTCCTGCAC 37639

Qy   2273 aagaaaacagaatttgtagttcacggtacaagtcttctttcaggactgttcaggtg 2332
          ||||||||| ||||||| || ||| ||||||||  | |||||||| |||| ||||
Db  37638 ATGGAGACAGGAAACTGTGTGGGTCACAGTCTCAGTCTCTTTTGGACCGTTTCAGACAACAAGCG 37579

Qy   2333 aagagaaatcaagaatttggggtttcacacagaaaaagagtggtatgtcaaatg 2392
          ||||||||||||||| ||| ||| |||| ||||||||| ||| | ||| ||||
Db  37579 ... 
```

```
RESULT  13

AC003083/c    AC003083    136823 bp    DNA    PRI    21-DEC-1999
LOCUS         AC003083    136823 bp    DNA    PRI    21-DEC-1999
DEFINITION    Homo sapiens BAC clone CTB-60N22 from 7q21, complete sequence.
ACCESSION     AC003083
VERSION       AC003083.1  GI:3947435
KEYWORDS      HTG.
SOURCE        human.
  ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 136823)
  AUTHORS     Waterston,R.
  TITLE       Direct Submission
  JOURNAL     Submitted (06-NOV-1997) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE     2  (bases 1 to 136823)
  AUTHORS     Kellen,J., Smith,A. and Gibson,A.
  TITLE       The sequence of Homo sapiens BAC clone CTB-60N22
  JOURNAL     Unpublished
REFERENCE     3  (bases 1 to 136823)
  AUTHORS     Waterston,R.
  TITLE       Direct Submission
  JOURNAL     Submitted (15-JAN-1998) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE     4  (bases 1 to 136823)
  AUTHORS     Waterston,R.
  TITLE       Direct Submission
  JOURNAL     Submitted (02-DEC-1998) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE     5  (bases 1 to 136823)
  AUTHORS     Waterston,R.
  TITLE       Direct Submission
  JOURNAL     Submitted (21-DEC-1999) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT       On Dec 2, 1998 this sequence version replaced gi:2588622.
              Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: saplens@watson.wustl.edu
                ------------- Summary Statistics
                Center project name: H_RG060N22
```

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GTB/CHR7/, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
Clone CTB-60N22 is from the first release of the human BAC library
CTB-978SK-B. The library contains cloned DNA from the male
fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl.
Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8
(1996). This clone is available from Research Genetics, Inc.
(http://www.resgen.com).
VECTOR: pBeloBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is CTB-135C18, 200 bp overlap; The
clone sequenced to the right is CTB-60P12, 200 base overlap.
Actual start of this clone is at base position 81036 of CTB-135C18;
actual end is at 36622 of CTB-60P12.

```
FEATURES             Location/Qualifiers
     source          1..136823
                     /organism="Homo sapiens"
                     /db_xref=taxon:9606"
                     /chromosome="7"
                     /map="7q21"
                     /clone="CTB-60N22"
                     /clone_lib="CTB-978SK-B"
     misc_feature    1160..2552
                     /note="CpG_island (%GC=66.5, o/e=0.90, #CpGs=140)"
     repeat_region   2977..3119
                     /rpt_family="L2"
     repeat_region   3498..3681
                     /rpt_family="MIR"
     repeat_region   4557..4608
                     /rpt_family="MIR"
     repeat_region   4609..4908
                     /rpt_family="Alu"
     repeat_region   4909..5002
                     /rpt_family="Alu"
     repeat_region   5495..5658
                     /rpt_family="MIR"
     repeat_region   5659..5737
                     /rpt_family="Alu"
     repeat_region   6015..6103
                     /rpt_family="L1"
     repeat_region   6210..6304
                     /rpt_family="L1"
     repeat_region   6318..6551
                     /rpt_family="L1"
     repeat_region   6552..6764
                     /rpt_family="L1"
     repeat_region   6768..7030
                     /rpt_family="MaLR"
     repeat_region   7272..7610
                     /rpt_family="Retroviral"
```

```
repeat_region   7403..7628
                /rpt_family="MER4-group"
repeat_region   7660..7898
                /rpt_family="Retroviral"
repeat_region   8013..8755
                /rpt_family="Retroviral"
repeat_region   8777..9094
                /rpt_family="L1"
repeat_region   9389..10058
                /rpt_family="Alu"
repeat_region   10069..10231
                /rpt_family="MER21_g"
repeat_region   10626..11262
                /rpt_family="L1"
repeat_region   11263..11827
                /rpt_family="L1"
repeat_region   11828..12240
                /rpt_family="Retroviral"
repeat_region   12241..12532
                /rpt_family="L1"
repeat_region   12533..12548
                /rpt_family="Alu"
repeat_region   12549..12849
                /rpt_family="L1"
repeat_region   12850..13189
                /rpt_family="Alu"
repeat_region   13205..13372
                /rpt_family="L1"
repeat_region   13376..13881
                /rpt_family="Alu"
repeat_region   13882..14120
                /rpt_family="L1"
repeat_region   14084..14283
                /rpt_family="L1"
repeat_region   14286..14760
                /rpt_family="L1"
repeat_region   14765..15133
                /rpt_family="L1"
repeat_region   15316..15669
                /rpt_family="MaLR"
repeat_region   17202..17287
                /rpt_family="L1"
repeat_region   17818..18025
                /rpt_family="L1"
repeat_region   18030..18105
                /rpt_family="L1"
repeat_region   18115..18424
                /rpt_family="Alu"
repeat_region   18791..18835
                /rpt_family="L1"
repeat_region   18888..19165
                /rpt_family="(CATA)n"
repeat_region   19778..20821
                /rpt_family="Alu"
repeat_region   20875..20991
                /rpt_family="Retroviral"
repeat_region   21983..42089
                /rpt_family="MIR"
gene            <21983..42089
                /gene="WUGSC:H_RG060N22.1"
CDS             join(<21983..22046,22044,24028..24134,28390..28457,30361,30485,
                31216..31389,34475..34584,36624..36705,41528..41608,
                41977..42089)
                /gene="WUGSC:H_RG060N22.1"
                /note="mitochondrial carrier protein-like; similar to
                O09461 (PID:g2479990); H_RG060N22.1"
                /codon_start=1
                /evidence=not_experimental
                /protein_id="AAC82534.1"
                /db_xref="GI:2588623"
                /translation="IVTPLDVVKIRLAQNNPLPKGKCFVYSNGLMDHLCVCEEGGNK
                LWYKKPGNPQGTLDAFFKIIRNEGIKSIWSGLPPTLWAVPATVIYFCYDQLSALLR
                SKLGENETCIPIVAGIVARFGAVTVISPLELIRTKMQSKKFSYVELHRVSKKVSEDG
                WISLWRGWAPTVLRDVPFSAMYWNYELLKWLCEKSGLYEPTFMINFTSGALSGSFA
                AVAPLPFDVVVKTQKQTOLWTYESHKISMPLHMSTWIIMKNIVARKNGFSGLFSGLIPRL"
```

```
repeat_region   23159..23457
                /rpt_family="Alu"
repeat_region   23507..23784
                /rpt_family="Alu"
repeat_region   24328..25740
                /rpt_family="L1"
repeat_region   25741..26036
                /rpt_family="Alu"
repeat_region   26037..26232
                /rpt_family="Alu"
                /rpt_family="L1"
```

```
                IKIAPACAIMISTYEFGKAFFQKQNVRROQY"
```

```
Query Match             20.4%;  Score 554.6;  DB 9;  Length 136823;
Best Local Similarity   99.3%;  Pred. No. 3.6e-95;
Matches   557;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy     5 cggcacgagctctctgaggctgcgccaagacctgaagcgcggccgacccggagggtct 64
         ||||| |||||||||||||| |||||||||||| ||||||||||||||||||||| |||
Db  2118 CCGCGCCAAACTCTCTGAGGGCTCCGCCAAGAACCTGAAGCGCGGCCGACCCGGGTCT 2059

Qy    65 gagactgagagagcaacgcagatggaagcgcgggggggtgagaggagacaacactgcaggggcc 124
Db  2058 GAGACTGAGAGAGCAACGCAGATGGAGGCGCGGGGGGGTGAGAGGCGGGAAACACAAACCTGCAGGGCC 1999

Qy   125 agagcgagggcgcgagagaaggagcgcgcgtgaggggggggcgcgcagcgcgcgaggagcgca 184
Db  1998 AGAGCGAGGGCGCGCGAGAAGGAGCGCGCGGCGGCGTGAGGGGGGGCGCGCGAGCGCGAGGGCA 1939

Qy   185 ggcacgaggggcggagcggcggaggcgggggcgcacggcggcggcgtgaggagggggggcg 244
Db  1938 GGCACGAGGGGCGGAGCGGCGGAGGCGGGGGCGCACGGCGGCGTGAGGAGGGGGGGCG 1879

Qy   245 cgtatggcggcgcgcggcggcgtgacggggtttcaaatcttcaaaatcttcaaaacgccgcagcccactcg 304
Db  1878 CGTATCGGCGCGCGGCGGCGTGACGGGTTTTCAAAATCTTCAAAATCTTCAACACGCGCCACTCG 1819

Qy   305 tttgtgctttgcgacctctcctccgccgcctggagcccggatccggcccgggaaaaccga 364
Db  1818 TTTGTGCTTTGCGCCCTTCCTCCTCCGCCGCCTGGAGCCCGGATCCGGCCCGGGAAACCCGA 1759

Qy   365 cctgcagacgcgggtacctctactgcgtagaggaggccgtagctgggtgggggaggggaggggagggggc 424
Db  1758 CCTGCAGACGCGGGTACCTCTACTGCGTAGAGGCCGTAGCTGGGTGGGGGAGGGAGGGAGGGGGC 1699

Qy   425 cgtcctgtcaaacggccgggggaaagccgtgctttctgggggctcccggggtcgcgacacttct 484
Db  1698 CGTCCTGTCAAACGGCCGGGGGAAGCCCGTGCTTTCGCGGGCTCCCGGGTCGCGACACTTTCT 1639

Qy   485 ccggaacccagacatgtaggtgccggggcgactgcaactccggagccatgaggaatccac 544
Db  1638 CCGGACCCAGACCATGTAGGTGCCGGGGCGACTGCCATGAACTCCGGAGCCATGAGGATCCAC 1579

Qy   545 agtaaaggaacattccagggt 565
         |||||||||||||||||||||
Db  1578 AGTAAAGGACATTTCCAGGGT 1558
```

```
RESULT  14
AX070167
LOCUS       AX070167               685 bp    DNA    PAT    25-JAN-2001
DEFINITION  Sequence 639 from Patent WO01102568.
ACCESSION   AX070167
VERSION     AX070167.1  GI:12579952
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 685)
AUTHORS     Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
            Kassan,A., Reinhard,C., Randazzo,F., Kennedy,C.C., Pot,D.,
            Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
            Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
```

```
TITLE     Human genes and gene expression products
JOURNAL   Patent: WO 0102568-A 639 11-JAN-2001;
          CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES             Location/Qualifiers
     source          1..685
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1..685
                     /note="n = A,T,C or G"
BASE COUNT      238 a      127 c      125 g      187 t        8 others
ORIGIN

Query Match         18.9%;   Score 513.8;  DB 6;  length 685;
Best Local Similarity  94.9%;  Pred. No. 1.6e-87;
Matches 560;  Conservative   0;  Mismatches  24;  Indels   6;  Gaps  3;

Qy   558 tccaggtggaatccaagtcaaaaatgaaaaaacagaccatctctgaaatcctgaaa 617
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TCCAGGTGGAATCCAAGTCAAAAATGAAAAAAACACAGACCATCTCTGAAATCCTGAAAA 60

Qy   618 ctgataacaggccagaaaaatccaaatgtaaggccacctttggggaagatattttaccttg 677
         |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    61 CTGATAACAGGCCAGAAAAATCCAAATGTAAGGCCACCTTTGGGGAAAAGTATTTTACCTTG 120

Qy   678 acttacccttctgtcacacatctgaaaaacttcaaaaggacattaaggatctgggagggc 737
         ||||||||||||||||| ||  |||||||||||||||||||||||||||||||||| |||
Db   121 ACTTACCTTCTGTCACCATATCTGAAAAACTTCAAAAGGACATTAAGGATCTGGGACGGC 180

Qy   738 gagttgaagaattctccagcaaagatatcagttatcttattcaaataaagaaggaagcta 797
         |||||||||||| || ||||||||||||| ||||||||||||||||||  ||||||| |||
Db   181 GAGTTGAAGAATTTCTCAGCAAAGATATATCAGTTATCTTATTCAAATAAGGAAGGCTA 240

Qy   798 aaattgcacacaaaccttgggtcgaattctccctgtaccaaagtccagaaatctgcatatactg 857
         |||||||||| |||||||||||||| |||| |||||| |||||||||||||||||||||||||||
Db   241 AAATTGCACAAACCTTGGGTCGGAATTTCTCCCTGTTACCAAAGTCCAGAAATCTGCATATACTG 300

Qy   858 cagaaacacttcacctcatcccagccatgatgaagttcattttaagtcaccacacacag 917
         ||||| ||| |||||||||||||||||||||| |||||||| |||||||||| |||||||
Db   301 CAGAAAACCACTTCACCTCATCCCAGCCATGATGGAAGTTCATTTAAGTCACCAGACACAG 360

Qy   918 tgtgtttaagcagagggaaaatattattagttgaaaaagctatccaaggaccatgatttcatc 977
         ||||||||||||||| ||||||||||||||||||||||||||| ||||||||||||| | | ||
Db   361 TGTGTTTAAGCAGAGGAAAATATTATTAGTTGAAAAAGCTATCAAGGACCATGATTTTATTC 420

Qy   978 cttcaaatagtatattcaaatgccttgtcatggggagtaaaaaatccttcatattgatg 1037
         ||||||||||||||||||| ||||||||||||| |||||||||||| ||||||||||||
Db   421 CTTCAAATAGTATATTCAAAATGCCTTGTCATGGGAGTAAAAAATTCTTCATATTGATG 480

Qy  1038 acaattagtactacattgaacaaaagagaaataaaagagttgtattattaccacaaaaatcaagta 1097
         ||| |||| |||||||||||||| ||||  ||  |||| || | |||||||||||| |||| |||
Db   481 ACATTAGATACTACATTGAACANAAGAAAAGAGNT--GATTTACTCACGAA--TCANGA 536

Qy  1098 cttcagtaagagatggggcaaaagagttggtagtggttgcacaaaaaca 1147
         ||||| |||| || |||||  |||| |||| |||||||| |||||||| |
Db   537 CTTCATTANNAGANTGGGG--CAAAAAGTTCTAGTGGGTCCCCATAAAACCA 584

RESULT  15
LOCUS       AC074175   177899 bp    DNA      HTG      15-JUL-2000
DEFINITION  Mus musculus chromosome 5 clone RP23-348B11 strain C57BL6/J,
            WORKING DRAFT SEQUENCE, 15 unordered pieces.
ACCESSION   AC074175
VERSION     AC074175.1  GI:9211286
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 177899)
  AUTHORS   Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
```

```
            Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,O.L., Maduro,V.B.,
            Mastrian,S.D., McCloskey,J.C., Ojodu,M.A., Pearson,R.,
            Stantipop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D. and Green,E.D.
  TITLE     NISC Mouse Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177899)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUL-2000) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     --------------- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc_mouse@nhgri.nih.gov
            --------------- Project Information
            Center project name: vc
            Center clone name: 348B11
            --------------- Summary Statistics
            Sequencing vector: plasmid; n/a; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 171084 bases at least Q40
            Consensus quality: 172951 bases at least Q30
            Consensus quality: 174122 bases at least Q20
            Insert size: 187000; agarose-fp
            Insert size: 202000; pulse-field-gel
            Insert size: 176499; sum-of-contigs
            Quality coverage: 4.87x in Q20 bases; agarose-fp
            Quality coverage: 4.50x in Q20 bases; pulse-field-gel
            Quality coverage: 5.16x in Q20 bases; sum-of-contigs
            -------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            -------------------------------------------------
            *       1     2463: contig of 2463 bp in length
            *    2464     2553: gap of unknown length
            *    2564     4767: contig of 2204 bp in length
            *    4768     4867: gap of unknown length
            *    4868     8397: contig of 3530 bp in length
            *    8398     8497: gap of unknown length
            *    8498    13167: contig of 4670 bp in length
            *   13168    13267: gap of unknown length
            *   13268    17790: contig of 4523 bp in length
            *   17791    17890: gap of unknown length
            *   17891    23556: contig of 5666 bp in length
            *   23557    23656: gap of unknown length
            *   23657    32999: contig of 9343 bp in length
            *   33000    33099: gap of unknown length
            *   33100    44567: contig of 11468 bp in length
            *   44568    44667: gap of unknown length
            *   44668    56459: contig of 11792 bp in length
            *   56460    56559: gap of unknown length
            *   56560    72129: contig of 15570 bp in length
            *   72130    72229: gap of unknown length
            *   72230    87873: contig of 15644 bp in length
            *   87874    87973: gap of unknown length
            *   87974   108836: contig of 20863 bp in length
            *  108837   108936: gap of unknown length
            *  108937   130636: contig of 21700 bp in length
            *  130637   130736: gap of unknown length
            *  130737   151003: contig of 20267 bp in length
            *  151004   151103: gap of unknown length
            *  151104   177899: contig of 26796 bp in length
FEATURES             Location/Qualifiers
     source          1..177899
```

```
misc_feature    /organism="Mus musculus"
                /strain="C57BL6/J"
                /db_xref="taxon:10090"
                /chromosome="5"
                /clone="RP23-348B11"
                /clone_lib="RPCI mouse BAC library 23"
                1..2463

misc_feature    /note="assembly_fragment"
                2564..4767
misc_feature    /note="assembly_fragment"
                4868..8397
                vector_side:left"
                clone_end:T7
misc_feature    /note="assembly_fragment"
                8498..13167
misc_feature    /note="assembly_fragment"
                13268..17790
misc_feature    /note="assembly_fragment"
                17891..23556
                clone_end:SP6
                vector_side:right"
misc_feature    /note="assembly_fragment"
                23657..32999
misc_feature    /note="assembly_fragment"
                33100..44567
misc_feature    /note="assembly_fragment"
                44668..56459
misc_feature    /note="assembly_fragment"
                5560..72129
misc_feature    /note="assembly_fragment"
                72230..87873
misc_feature    /note="assembly_fragment"
                87974..108836
misc_feature    /note="assembly_fragment"
                108937..130636
misc_feature    /note="assembly_fragment"
                130737..151003
misc_feature    /note="assembly_fragment"
                151104..177899
                /note="assembly_fragment"

BASE COUNT      53941 a   34927 c   35777 g   51822 t   1432 others
ORIGIN

Query Match          15.6%;    Score 423.6;  DB 2;   Length 177899;
Best Local Similarity   64.3%;  Pred. No. 2.5e-70;
Matches  779;  Conservative   0;  Mismatches  394;  Indels  39;  Gaps  8;
```

```
Qy  1502 aagaataaaatacagtgtttggatcccttctcctgttcctgcaaagtgtcctgaaaaga 1561
         |||| || ||||||| | |||||||| ||| | ||| |||||||||| |||||||||
Db 98980 AAAGAAATAAGATACAGTGTCGGGTCCCTTCTCTCGTTCTGCAAATGTCCTGAAAAACA 99039

Qy  1562 ctgaacaaaaggaaaagtggaattgcaacaatattctcagaaaagaaattgccagaagatg 1621
         |||| |||  ||| || |||||||||||| | ||||||||||||||| |||||| ||
Db 99040 CTGCACCGAAGGAGGAAGCCACTCTTGGAGCCCAATTCCAGAAGGACGTGGGCGAAAGCA 99099

Qy  1622 atacaacagtgaaggagcagaattcctgtataaagagaccaggaaactgaaaaaaagc 1681
         ||  ||||| | ||||||||||||||| ||| || |||||||| |||||| ||
Db 99100 GTGGGCACCTGCTCAAGCCGAAACTCTCAGTGATGAAGAAACCCAGAAAACTTGAAGAAAAGC 99159

Qy  1682 tcctgtttatttcagagccccatcccccccaccctccaaatgaattgagagggcttaatgaga 1741
         || | || |||||||||||||| ||||| ||| |||||||| |||| ||||||||
Db 99160 ATGGGTTTCTTCAGAGCCCCACACCTATTCTTCTTCAGCTGATTGAAAAGGATGTGA---- 99214

Qy  1742 aaatgagtaataaaatgttccatgttaagtacagctgaagatgacataagacagaattta 1801
         ||||||| |||||| |||||||||| |||||| |||||||| ||||||||||||||
Db 99215 ----CAGGAAAACAGTGAGCAGTGTTCAATGCAAGTGAGCCTGACCAGAGCAGGAGTATG 99270

Qy  1802 cacagctacctctacataaacaacacaggaaatgcatttcttgacatttccgaacacacat 1861
         ||  |||||||||||||||||| |||||||||| |||||||| | |
Db 99271 CACAGCTACCTCTCGCGTGACAGCACACCGGAA-----------CACCAATGACCG 99315

Qy  1862 taagtgaaaatgacttagaagaactaaggtagatcactataaatgtaacatacacaggcat 1921
```

```
Db 99316 AGGGTAGAAAATGAC---GGAGAACAGAGAGTAGAACCCTGCTTCCAGGGGTGTCGGAGTCAT 99372
         || |||||| ||| ||    | ||| ||||||| ||| ||| |||||||||||||| ||

Qy  1922 ctgtacaatgttcctgatttcagtacagatataagtggatctcaaccaaacagagagtcag 1981
         ||| ||||||| ||||| | |||| ||||||| ||||| ||||||| ||||| ||||||| |
Db 99373 GTGGACAAGTTTCTCATCTCAGTACACAGATACAGAAAGTAATCTACCCCAACACAGTTAGCTGCAG 99432

Qy  1982 atactgtgtcttttccagcaaaggacctcaaggaaaaggaccttcattcattcaattattctc 2041
         |||| | || ||||||||||||||||| ||||||||||| ||||||| ||| ||| ||||
Db 99433 ACATTACACAGCTTCACTCTCAGTACACTACAGAAAGTGCAAGAGATCTGCAAGAATCTAACTAACTACAA 99492

Qy  2042 atgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaagg 2101
         || |||||||||| | ||||||||||||||||||| |||| | |||| |||||||||||
Db 99493 ATGCTTCTGAACCCAGTGCTGAAATTGGATAAAGCAGAAAATACATCAGACACAGAAAAGTAATCTAC 99552

Qy  2102 ctccattccaatactcctcctgaggaacccaatgaatgtgacttcaagatatggatag 2161
         || |||||||| ||| ||||| ||||| |||||||||| ||||||||||| ||||
Db 99553 CTCCACCCTGTAGTCCCC----AGGAGCCTCAГGAGTGTGACACTGAGACAACACAGGAGAATT 99609

Qy  2162 tacccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaaataqaaaagaaa 2221
         |||| ||||||||||||||| | ||||| |||||||| ||| ||||||| ||
Db 99610 TGCCTTGGTGGTAAAAATCCAGCGGAAAGTGAGAAATGCTATTTAGGGGCAACAG---AAAGCAA 99666

Qy  2222 atcctggaaccaaatgctgaattctgataaa---agaactgaatttattacacaagaagaa 2278
         | |||| |||||  ||||| || || ||| |     | |||| ||| |||| ||||||||
Db 99667 ATGCTGAACCGAGTGCTGAAATTGGATAAAGCAGAAAATACATCAGACAACTGAGACTTCCTGCACATGAAG 99726

Qy  2279 acagaatttgtagttcaccggtacagtctttactagacttgtttcagagaggagaga 2338
         ||| || |||||| ||| || ||| ||||| |||| || ||||| |||
Db 99727 ACAGAACTTGTGTGGTCACCAGTACACTGCTTCTTGGACTTTTGTTCCAGAGGGGAGAGA 99786

Qy  2339 aatcagaagaatttttgggtttcacaagctacacacagaaagagtgggtatatgcaatgtttttag 2398
         |||||||||| ||||| ||||||||||| ||| ||| ||||||||| || | |||
Db 99787 AATCAGAATTTTTGGGGTTTTACAGGCTACAGTAGGAAATATGTGAATGTTTTAG 99846

Qy  2399 atatttgggaaggaaaaaatccagataaatctgttaacacgttttttctcgtccccttcaa 2458
         ||||||||| |||||||| || ||| ||||| | || |||||| ||||||
Db 99847 ATATTTGGGAGGAAGAAAAATTCAAGTACCCCTGTTTGTCAAACGTGTTTTTTCTTGTCCCCCCTTCAA 99906

Qy  2459 cttctacattactggctcttttagaaattcaaaaaaatgcatactttcagaagtgataagga 2518
         |||| ||||||| | | || |||||||||| ||||||| | ||| | || ||
Db 99907 CTTCTGCATTCGTTCGGGAATTTAGTATTTTTAGTATTTTAAAATGTGCTTTTCAGAAGTGACACTGTATATTCC 99966

Qy  2519 tcatattccttgaaatttttataaatatgtatggaaattcctaggaattcttttttaccagctt 2578
         || | |||| ||||  ||||| || |||||||||||||||||||| || || ||
Db 99967 TGAAGTTTTATATAAAATACATATATATATAATTAATTAGTGGTGGTGGGGGTTTGGTTGTTTTCGGCCAACT 100026

Qy  2579 tgtttcagaccaaatgtaaataattattaa--aaataaatatattgcaatttctacagaatt 2636
         ||||| ||| |||||| |||| |||||       |||||||||| |||||| |||||||||
Db 100027 TTTTTATGGAGCGCATCTGTAAATATAATTAATTAGGTGGTTGTGTTTGTAATTTTTTTACAGGAATT 100086

Qy  2637 gaataccتgtaaagaaaattacagaataaacttgtgactggtcttgttttacattaaa 2696
         |||| | | || |||| |||||| | ||
Db 100087 GACTAGCTGTTACA-AATTATACTAAAAAAACTTAAAGACTTATTTTTTATATTCCTGAGAA 100145

Qy  2697 aaaaaaaaaa 2708
         ||||||||||
Db 100146 TAAAAATGAATA 100157
```

```
Search completed: December 27, 2001, 20:56:56
Job time: 14571 sec
```